

**DIVERSITY IN THE BOREAL FOREST OF ALASKA: DISTRIBUTION AND  
IMPACTS ON ECOSYSTEM SERVICES**

**A**

**THESIS**

**Presented to the faculty  
of the University of Alaska Fairbanks  
in Partial Fulfillment of the Requirements  
for the Degree of**

**DOCTOR OF PHILOSOPHY**

**By**

**Brian D. Young, B.S., M.S.**

**Fairbanks, Alaska**

**December 2012**

UMI Number: 3537837

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.

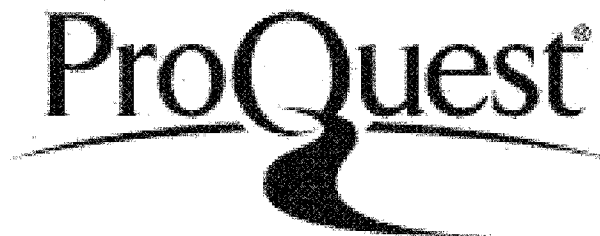


UMI 3537837

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code.



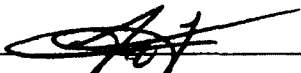

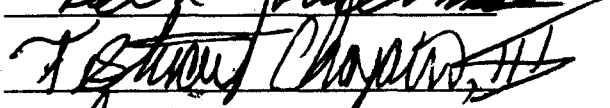

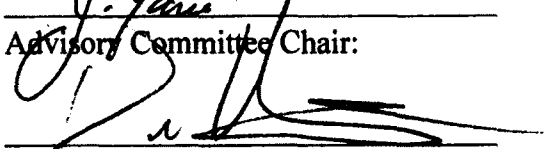
ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

DIVERSITY IN THE BOREAL FOREST OF ALASKA: DISTRIBUTION AND  
IMPACTS ON ECOSYSTEM SERVICES


By

Brian D. Young, M.S.

RECOMMENDED:

  
David Verbyla  
  
Zoltan Papp  
  
Robert Chapin, III  
  
J. Gair  
Advisory Committee Chair:  
  
Chair, Department of Forest Science

APPROVED:

  
S. D. Sp  
Dean, Natural Resources and Agricultural Sciences

  
John C. Eichberger  
Dean of the Graduate School

11/29/2012  
Date

## ABSTRACT

Within the forest management community, diversity is often considered as simply a list of species present at a location. In this study, diversity refers to species richness and evenness and takes into account vegetation structure (i.e. size, density, and complexity) that characterize a given forest ecosystem and can typically be measured using existing forest inventories. Within interior Alaska the largest forest inventories are the Cooperative Alaska Forest Inventory and the Wainwright Forest Inventory. The limited distribution of these inventories constrains the predictions that can be made. In this thesis, I examine forest diversity in three distinct frameworks; Recruitment, Patterns, and Production. In Chapter 1, I explore forest management decisions that may shape forest diversity and its role and impacts in the boreal forest. In Chapter 2, I evaluate and map the relationships between recruitment and species and tree size diversity using a geospatial approach. My results show a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in their effects on recruitment. In Chapter 3, I modeled and mapped current and possible future forest diversity patterns within the boreal forest of Alaska using machine learning. The results indicate that the geographic patterns of the two diversity measures differ greatly for both current conditions and future scenarios and that these are more strongly influenced by human impacts than by ecological factors. In Chapter 4, I developed a method for mapping and predicting forest biomass for the boreal forest of interior Alaska using three different machine-learning techniques. I developed first time high resolution prediction maps at a 1km<sup>2</sup> pixel size for aboveground woody biomass. My results indicate that the geographic patterns of biomass are strongly influenced by the tree size class diversity of a given stand. Finally, in Chapter 5, I argue that the methods and results developed for this dissertation can aid in our understanding of forest ecology and forest management decisions within the boreal region.

## TABLE OF CONTENTS

SIGNATURE PAGE .....	i
TITLE PAGE .....	ii
ABSTRACT.....	iii
TABLE OF CONTENTS.....	iv
LIST OF FIGURES .....	ix
LIST OF TABLES.....	xiii
LIST OF APPENDICES.....	xv
PREFACE .....	xviii
CHAPTER 1: ECOLOGY AND MANAGEMENT OF THE BOREAL FOREST OF INTERIOR ALASKA.....	1
1.1 INTRODUCTION.....	1
1.2 REFERENCES.....	7
CHAPTER 2: EFFECTS OF SPECIES AND TREE SIZE DIVERSITY ON RECRUITMENT IN THE ALASKAN BOREAL FOREST: A GEOSPATIAL APPROACH .....	16
2.1 ABSTRACT.....	16
2.2 INTRODUCTION.....	17
2.3 DATA AND METHODS.....	19

2.3.1 DATA .....	19
2.3.2 METHODS .....	23
2.4 RESULTS.....	25
2.4.1 ATTRIBUTES OF THE TWO FOREST TYPES.....	25
2.4.2 FACTOR IDENTIFICATION AND PARAMETER ESTIMATION .....	26
2.4.3 EFFECTS OF TREE SIZE AND SPECIES DIVERSITY ON RECRUITMENT .....	27
2.4.3.1 TREE SIZE DIVERSITY .....	27
2.4.3.2 TREE SPECIES DIVERSITY .....	28
2.5 DISCUSSION AND CONCLUSION .....	29
2.6 REFERENCES.....	32
 CHAPTER 3: MODELING AND MAPPING FOREST DIVERSITY OF INTERIOR ALASKA AT 1-KM <sup>2</sup> RESOLUTION FOR CURRENT AND POSSIBLE FUTURE CLIMATE CONDITIONS .....	 118
3.1 ABSTRACT .....	118
3.2 INTRODUCTION.....	118
3.3 METHODS AND MATERIALS .....	122
3.3.1 DIVERSITY DATA .....	122
3.3.2 ENVIRONMENTAL FACTORS.....	122

3.3.3 THE CALIBRATION AND VALIDATION DATASETS .....	125
3.3.4 STATISTICAL METHODS.....	125
3.3.5 PREDICTIVE MAPS .....	127
3.4 RESULTS.....	127
3.4.1 PREDICTOR VARIABLE CORRELATIONS .....	127
3.4.2 VARIABLE SELECTION AND IMPORTANCE.....	127
3.4.3 SPATIAL DEPENDENCY OF TREE SIZE-CLASS AND SPECIES DIVERSITY .....	128
3.4.4 RANDOM FOREST ANALYSIS MODEL ASSESSMENT FOR TREE SIZE- CLASS AND TREE SPECIES DIVERSITY .....	128
3.4.5 CURRENT PREDICTED DIVERSITY PATTERNS .....	130
3.4.5.1 TREE SIZE-CLASS DIVERSITY .....	130
3.4.5.2 TREE SPECIES DIVERSITY .....	130
3.4.6 PREDICTED FUTURE DIVERSITY PATTERNS ASSUMING NO MAJOR DISTURBANCE EVENTS.....	131
3.4.6.1 TREE SIZE-CLASS DIVERSITY .....	131
3.4.6.2 TREE SPECIES DIVERSITY.....	132
3.5 DISCUSSION .....	133
3.6 REFERENCES.....	138

<b>CHAPTER 4: MAPPING AND PREDICTING ABOVEGROUND BIOMASS OF TREES USING FOREST INVENTORY DATA AND PUBLIC ENVIRONMENTAL VARIABLES WITHIN THE ALASKAN BOREAL FOREST .....</b>	<b>206</b>
<b>4.1 ABSTRACT .....</b>	<b>206</b>
<b>4.2 INTRODUCTION.....</b>	<b>206</b>
<b>4.3 METHODS AND MATERIALS .....</b>	<b>209</b>
<b>4.3.1 BIOMASS DATA .....</b>	<b>209</b>
<b>4.3.2 ENVIRONMENTAL FACTORS.....</b>	<b>210</b>
<b>4.3.3 THE CALIBRATION AND VALIDATION DATASETS .....</b>	<b>211</b>
<b>4.3.4 STATISTICAL METHODS.....</b>	<b>212</b>
<b>4.3.5 PREDICTIVE MAPS .....</b>	<b>214</b>
<b>4.3. RESULTS.....</b>	<b>215</b>
<b>4.3.1 VARIABLE SELECTION AND IMPORTANCE.....</b>	<b>215</b>
<b>4.3.2 RANDOM FOREST ANALYSIS MODEL ASSESSMENT .....</b>	<b>215</b>
<b>4.3.3 INFLUENCE OF THE ENVIRONMENTAL FACTORS ON ABOVEGROUND FOREST BIOMASS.....</b>	<b>216</b>
<b>4.3.4 REGRESSION TREE ANALYSIS MODEL ASSESSMENT .....</b>	<b>217</b>
<b>4.3.5 SPATIAL DEPENDENCY OF ABOVEGROUND FOREST BIOMASS ....</b>	<b>217</b>
<b>4.3.6 PREDICTED ABOVEGROUND FOREST BIOMASS PATTERNS .....</b>	<b>217</b>



4.4. DISCUSSION .....	218
4.6 REFERENCES.....	221
CHAPTER 5: CONCLUSION .....	253
5.1 REFERENCES.....	258

## LIST OF FIGURES

Figure 1.1: Geographic distribution of the 704 Sample Plots (in triangles) within the Alaskan boreal forest (Ruefenacht et al., 2008).....	14
Figure 1.2: Distribution of species by volume of total growing stock (A) and stand diameter breast height (dbh) size classes (B) in the forests of interior Alaska. Sawtimber: 22.9cm DBH or greater, Poletimber: 15cm to 22.8cm DBH, Reproduction: 2.5cm to 14.9cm DBH (Malone et al., 2009).....	15
Figure 2.1: Geographic distribution of the 438 Sample Plots (in dots) and the mixed poplar/birch and mixed spruce forest types within the State of Alaska (Ruefenacht et al., 2008). ....	49
Figure 2.2: Maps of expected sapling recruitment over a 10-year period for the mixed poplar/birch (Figs A, B, and C) and the mixed spruce (Figs D, E, and F) forest types respectively predicted from high ( $H_d = 2.51$ ; Figs A and D), median ( $H_d = 1.89$ ; Figs B and E), and low ( $H_d = 1.27$ ; Figs C and F) tree size diversity within the State of Alaska.	50
Figure 2.3: Average basal area ( $m^2 ha^{-1}$ ) of trees on sites with low tree size diversity ( $H_d < 1.3$ ; dark green bars) and high tree size diversity ( $H_d > 2.3$ ; light green bars) within the 20 different size classes (see text for further description) from mixed poplar/birch stands (a) and mixed spruce stands (b). ....	51
Figure 2.4: Maps of expected sapling recruitment over a 10-year period for the mixed poplar/birch (Figs A, B, and C) and the mixed spruce (Figs D, E, and F) forest types respectively predicted from high ( $H_s = 1.23$ ; Figures A and D), median ( $H_s = 0.59$ ; Figures B and E), and low ( $H_s = 0.00$ ; Figures C and F) species diversity within the State of Alaska. ....	52

Figure 3.1: Geographic distribution of the 704 Permanent Sample Plots (in triangles) within the Alaskan boreal forest (Ruefenacht et al., 2008).....	155
Figure 3.2: Hierarchical clustering using squared Spearman correlation ( $p^2$ ) of the environmental variable similarities.....	156
Figure 3.3: Application of the RFA model for tree size-class diversity ( $H_d$ ) to the calibration data set (Cal) and the validation data set (Val). ....	157
Figure 3.4: Application of the RFA model for tree species diversity ( $H_s$ ) to the calibration data set (Cal) and the validation data set (Val). ....	158
Figure 3.5: Partial plots representing the marginal effects of the twelve most influential variables from the RFA model on the estimates of tree size-class diversity ( $H_d$ ) while averaging out the effect of all the other variables. ....	159
Figure 3.6: Partial plots representing the marginal effects of the twelve most influential variables from the RFA model on the estimates of tree species diversity ( $H_s$ ) while averaging out the effect of all the other variables. ....	160
Figure 3.7: Current predicted tree size-class diversity ( $H_d$ ) in the Alaskan boreal forest using the RFA model. ....	161
Figure 3.8: Current predicted tree species diversity ( $H_s$ ) in the Alaskan boreal forest using the RFA model. ....	162
Figure 3.9: Predicted changes in tree size-class diversity ( $H_d$ ) from its current level (Figure 3.6) to the year 2050 for a null climate change scenario (A) and the IPCC A1B scenario (B). ....	163

Figure 3.10: Predicted changes in tree species diversity ( $H_s$ ) from its current level (Figure 3.7) to the year 2050 for a null climate change scenario (A) and the IPCC A1B scenario (B). .....164

Figure 4.1: Geographic distribution of the 704 Sample Plots (in triangles) within the Alaskan boreal forest (Ruefenacht et al., 2008).....236

Figure 4.2: Application of the RFA model for woody biomass to the calibration data set (Cal) and the validation data set (Val). .....237

Figure 4.3: Predicted and observed above ground full tree woody biomass (Mg/ha dry weight) for the 704 Sample Plots within the Alaskan boreal forest. The current model (black dots) indicates the model presented in this paper while the other predictions are from Blackard et al. (2008, red triangles). .....238

Figure 4.4: Partial dependence plots for the six most important variables from the random forest analysis (RFA; Table 4.3) predictions of aboveground forest woody biomass on tree size-class diversity ( $H_d$ ), Vegetation type (Veg, refer to Table 4.2), June precipitation ( $P_{06}$ ), Distance to communities (Dtc), Elevation (El), and Longitude (Y). Partial dependence is the predicted value of the response based on the value of one predictor variable after averaging out the effects of the other predictor variables in the model.....239

Figure 4.5: Conditional Inference tree for aboveground forest woody biomass using the six most important predictor variables as determined by the random forest analysis (see Figure 4.4, Table 4.3). The p-values at each of the nodes are from a Monte Carlo randomization test and in order for a split to occur the p-value must be  $<0.001$ . The box plots at the terminal nodes represent the distribution of the data within that branch of the tree. The shaded areas in the boxes represent the inner-quartile range while the horizontal lines represent the median values and, the whiskers extend to the most extreme data point that is no more than 1.5 times the inner-quartile range. ....240

**Figure 4.6: A map of the aboveground forest woody biomass (Mg/ha dry weight) derived from the random forest analysis model using the CAFI and WAIN forest inventory plot biomass values predicted across the landscape as a function of environmental predictors (Table 4.3). .....241**

## LIST OF TABLES

Table 2.1: Frequency of sampled tree species by forest type. ....	42
Table 2.2: Definition, mean, and standard deviation (S.D.) of plot variables used in the analysis.....	43
Table 2.3: Percentage contribution (%) to the overall goodness-of-fit and the level of significance (*P < 0.05; **P < 0.01; ***P < 0.001) for the explanatory variables of sapling recruitment within mixed poplar/birch and mixed spruce forest types. ....	44
Table 2.4: Difference between the geospatial and non-spatial models of sapling recruitment. ....	45
Table 2.5: Pearson correlations and their levels of significance (*, P < 0.05; **, P < 0.01; ***, P < 0.001) between tree size diversity ( $H_d$ ), species diversity ( $H_s$ ), and other explanatory variables for mixed poplar/birch and mixed spruce forest types. ....	46
Table 2.6: Spatial autocorrelation and its level of significance of linear recruitment model (Eq. 2.3) residuals for mixed poplar/birch and mixed spruce forests types in Alaska. ....	47
Table 3.1: Definition of variables used in the analysis.....	152
Table 3.2: Variable importance (ranking) in determining tree size-class diversity ( $H_d$ ) and tree species diversity ( $H_s$ ) using percent increase in mean standard error (%IncMSE) for ranking purposes. All variables listed in this table were determined to be significant based on the Z-score with an upper 95% confidence limit ( $ Z  \leq 1.65$ ) and were used in the development of the final random forest analysis models.....	153

Table 3.3: Spatial autocorrelation and its level of significance for Tree Size Class Diversity ( $H_d$ ) and Tree Species Diversity ( $H_s$ ) within the Alaskan boreal forest, and for the residuals of the random forest analysis (RFA) models used for predicting $H_d$ and $H_s$ from this dataset. ....	154
Table 4.1: Definition of variables used in the analysis.....	231
Table 4.2: Vegetation variables (Veg, from Table 4.1) used in this analysis were derived from the classification system developed by Fleming (1997). ....	233
Table 4.3: Variable importance (ranking) in determining aboveground forest woody biomass (biomass; Mg/ha dry weight) using percent increase in mean standard error (%IncMSE) for ranking purposes. The variables deemed important were used in the development of the final random forest analysis (RFA).....	234
Table 4.4: Spatial autocorrelation and its level of significance for aboveground forest woody biomass (biomass; Mg/ha dry weight) in the Alaskan boreal forest, and for the residuals of the random forest analysis (RFA) model used for predicting biomass from this dataset.....	235
Table 5.1: Open Access spatial data for the boreal forest of Alaska .....	261

## LIST OF APPENDICES

Appendix 2.1: R code used in the analysis for Chapter 2 for the poplar/birch forest type .....	53
Appendix 2.2: R code used in the Analysis for Chapter 2 for the mixed spruce forest type .....	59
Appendix 2.3: Mixed poplar/birch and mixed spruce forest types within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	66
Appendix 2.4: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from high ( $H_d = 2.51$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998.....	70
Appendix 2.5: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from median ( $H_d = 1.89$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	74
Appendix 2.6: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from low ( $H_d = 1.27$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998.....	78
Appendix 2.7: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from high ( $H_d = 2.51$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998.....	82
Appendix 2.8: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from median ( $H_d = 1.89$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998.....	86



Appendix 2.9: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from low ( $H_d = 1.27$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998.....	90
Appendix 2.10: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from high ( $H_s = 1.23$ ) tree species diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	94
Appendix 2.11: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from median ( $H_s = 0.59$ ) tree species diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	98
Appendix 2.12: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from low ( $H_s = 0.00$ ) tree species diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	102
Appendix 2.13: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from high ( $H_s = 1.23$ ) tree species diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998.....	106
Appendix 2.14: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from median ( $H_s = 0.59$ ) tree species diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998.....	110
Appendix 2.15: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from low ( $H_s = 0.00$ ) tree species diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998.....	114
Appendix 3.1: R code used in the analysis for Chapter 3.....	165
Appendix 3.2: Forest types of the boreal forest of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	177

Appendix 3.3: Current predicted tree species diversity ( $H_s$ ) in the Alaskan boreal forest Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	182
Appendix 3.4: Current predicted tree size-class diversity ( $H_d$ ) in the Alaskan boreal forest Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	186
Appendix 3.5: Predicted changes in tree species diversity ( $H_s$ ) from its current level to the year 2050 for a null climate change scenario Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	190
Appendix 3.6: Predicted changes in tree species diversity ( $H_s$ ) from its current level to the year 2050 for a IPCC A1B scenario Raster Dataset Metadata Standard Version FGDC- STD-001-1998 .....	184
Appendix 3.7: Predicted changes in tree size-class diversity ( $H_d$ ) from its current level to the year 2050 for a null climate change scenario Raster Dataset Metadata Standard Version FGDC-STD-001-1998.....	188
Appendix 3.8: Predicted changes in tree size-class diversity ( $H_d$ ) from its current level to the year 2050 for a IPCC A1B scenario Raster Dataset Metadata Standard Version FGDC-STD-001-1998 .....	192
Appendix 4.1: R code used in the Analysis of Chapter 4 .....	242
Appendix 4.2: Aboveground forest woody biomass (Mg/ha dry weight) within the boreal forest of Alaska Raster Dataset.....	236

## PREFACE

This dissertation could not have been completed without the help and extensive support from several organizations and individuals. Deep thanks go to my graduate committee of John Yarie, Terry Chapin, Faulk Huettmann, Dave Verbyla, and Josh Greenburg. Your hours of dedication are extremely appreciated. I received support for this work from the National Science Foundation, through its Integrative Graduate Education and Research Traineeship (IGERT) to the Resilience and Adaptation Program (RAP) at the University of Alaska Fairbanks, the Alaska EPSCoR NSF award #EPS-0701898 and, the State of Alaska Division of Forestry. I am also grateful to Ellen Weiser, and Gary Kofinas for their tireless efforts in making the RAP program even better. A special thanks also goes to Diane Wagner for her diligence and commitment to making me a better writer and scientist. I would like to thank my wife Sara for without her, none of this would have been possible! I would also like to thank my parents for always encouraging me to take the more difficult path. Lastly, something that I learned while serving as a Peace Corps Volunteer in Senegal which has kept me true to my path: Dank n' dank moy japp golo chi nigh!

## CHAPTER 1: ECOLOGY AND MANAGEMENT OF THE BOREAL FOREST OF INTERIOR ALASKA

### 1.1 INTRODUCTION

A difficult task that forest managers and decision makers face is the definition, allocation, and distribution of sustainable forest management (SFM) practices over time and across landscapes. These difficulties are compounded by the desire to meet present and future competing demands on forests while conserving important natural resources. While the term “SFM” has been widely accepted (Charron 2005), a clear definition has been elusive (Wang 2004) but, it can be generally thought of as the balancing of ecological, social, and economic values to meet society’s objectives over the long term (Sheppard 2005). However, what some in the forest management community consider a simple concept involves a very challenging task of developing and implementing management strategies, environmental commitments, and policies while applying adaptive management to account for emerging social needs and global trends. The complexity of this task lies not only in the challenge of integrating and comparing diverse social, ecological, and economic interests, but also in the lack of methodologies that allow us to quantify and compare the value of many forest ecosystem benefits. The valuation, quantification, and geographical location of forest ecosystem benefits allows us to identify appropriate management goals, anticipate social reactions, and deal with conflicts over forest lands (Bengston 1994).

While the issue of sustainable forests is not new, it has taken on new meaning and urgency in recent decades (Burton et al. 2003; Von Gadow et al. 2001). The roots of SFM extend back to at least 1346, when King Philippe of France decreed that forests are to be continuously maintained and kept in good condition (Forestry 2012). Later, in the United States, the very notion of sustained yield was at the heart of the conservation idea espoused by Chief Forester Gifford Pinchot early in the 20th century (Parnell 2012). This concept was later applied to the Multiple Use-Sustained Yield Act of 1960 (MUSYA) (16

U.S.C. Sec. 528-531 [1976]), which extended the sustained yield principles in the United States to not only cover timber, but also outdoor recreation, watersheds, and wildlife and fish resources. SFM has become more globally recognized because of changes in societal values towards sustainable development that were highlighted in the Brundtland report *Our Common Future* (Brundtland 1987).

For effective SFM implementation, an adaptive management (AM) framework can be applied (Foster et al. 2010; Holling 1978). AM involves four decision-making stages: planning, implementation, evaluation, and modification (Walters and Holling 1990) thus creating a feedback loop that can be closed via periodic monitoring and revision (Bormann et al. 2007). The linking of SFM and AM has created adaptive forest management (AFM) which aims to preserve and develop the functionality of forests as a prerequisite for fulfilling the future need for forest ecosystem services (Wagner 2004). AFM can perhaps best be initiated through the incorporation of predictive models, scenarios, and the use of open access data (Huettmann 2007; Walters 1986; Wollenberg et al. 2000). In the future, as new knowledge is gained and data are shared, the models can be updated so that management decisions are then adapted which in turn makes the process highly dynamic and dependent upon continuous research.

Permanent Sample Plots (PSPs) are an invaluable tool for resource managers, in part because they provide managers with a wide variety of data and are re-measured on a periodic basis. While a single measurement of a given PSP gives a snapshot in time of various attributes of interest at a given location, the periodic re-measurements of PSPs provide much more valuable long-term information on dynamic processes. A well-established system of PSPs can provide for the monitoring for the consequences of large-scale environmental changes across large areas over time (Bakker et al. 1996; Poso 2006; Stott 1947). The data within the PSPs, combined with remote sensing data from both satellite and aerial platforms can allow for the mapping of various forest attributes and dynamic processes at the landscape scale (see for instance Fassnacht et al. 2006; Iverson and Prasad 2001; McRoberts et al. 2008; Ruefenacht et al. 2008).

Within the boreal forest of Alaska two separate forest inventories use a PSP design, the Cooperative Alaska Forest Inventory (CAFI) and the Wainwright Forest Inventory (WAIN) (Malone et al. 2009; Ress, personal communication). These forest inventories are located primarily on well-stocked forested lands and together consist of over 704 PSPs. The CAFI plots are primarily located along the road system on Federal, State, Borough, and Native Corporation lands, while the WAIN plots are scattered across Military lands (Figure 1.1). These forest inventories contain the largest collection of field-gathered data on the forest conditions within boreal Alaska. While other forest inventories do exist within the boreal forest of Alaska, notably those conducted on Native Corporation Lands ([www.tananachiefs.org](http://www.tananachiefs.org)) and on Forest Classified State Lands ([forestry.alaska.gov](http://forestry.alaska.gov)), they did not utilize a PSP design. The Forest Inventory and Analysis (FIA) Program of the U.S. Forest Service which is directed to “make and keep current a comprehensive inventory and analysis of the present and prospective conditions of and requirements for the renewable resources of the forest and rangelands of the United States.” as mandated by the Forest and Rangeland Renewable Resources Research Act of 1978 (16 U.S.C. §1641) is only now being initiated (2011-2013) in Interior Alaska.

The boreal forest of Alaska is the predominate ecoregion within the state. This forest extends from the Kenai Peninsula to the foothills of the Brooks Range and from the Porcupine River near the Canadian border west to the Kuskokwim River valley (Figure 1.1). The vegetation within this forest type is comprised of a mosaic of stands of different ages and sizes (Fig 1.2). There are eight species currently present within this forest type consisting of white spruce (*Picea glauca* (Moench) Voss), black spruce (*Picea mariana* (Mill.) B.S.P.), tamarack (*Larix laricina* (DuRoi) K. Koch), Kenai birch (*Betula kenaica* W.H. Evans), Alaska birch (*Betula neoalaskana* Sarg.), quaking aspen (*Populus tremuloides* Michx.), balsam poplar (*Populus balsamifera* L. spp. *balsamifera*), and western black cottonwood (*Populus balsamifera* L. spp. *trichocarpa* (Torr. & Gray) (Vioreck and Little 2007).

Surprisingly, little is known about the role of diversity and ecosystem functioning within the boreal forest (Nadrowski et al. 2010). Although tree species diversity within the boreal forest is low (Burton et al. 2003), the effects of diversity may still positively influence productivity (Tilman et al. 2001). Diversity is often considered as simply a list of species present at a given location. However, this does not sufficiently describe the diversity of a forested stand. Structural diversity, resulting from recruitment of trees of different sizes into multilayered canopies, should also be taken into account. This characteristic, which can be approximated by the diversity of tree size, affects the amount of light and precipitation received by subordinate trees and understory plants (Anderson et al. 1969) and may thus influence the productivity of forest ecosystems.

Silvicultural treatments are often defined by target stand states defined by the distribution of trees by size class (Smith et al. 1997). Manipulating tree-size diversity is thus a practical tool for forest managers who strive for greater biodiversity and/or greater productivity (Varga et al. 2005). Previous studies dealing with tree-size diversity include Oren et al. (1987) and Lusk and Ortega (2003). Liang et al. (2007) considered both the effects of tree-species diversity and tree-size diversity on individual tree growth, mortality, and recruitment in the Pacific Northwest. While this study did show a positive effect of species and size class diversity on productivity, this relationship is not universal (Homeier et al. 2010; Lei et al. 2009; Varga et al. 2005; Vila et al. 2007; Vila et al. 2003). Therefore, much still needs to be learned about the effects of species and size class diversity on forest productivity.

Management decisions typically affect many forest attributes including tree diameter distribution (Buongiorno and Gillespie 2003; Lin et al. 1998; Schwartz et al. 2005; Shao et al. 2005), species composition (Fu et al. 2007; Schwartz et al. 2005), diversity (Eriksson and Hammer 2006), and the amount of litter and coarse woody debris (Alban et al. 1994; Duvall and Grigal 1999). Forest biomass is also affected through management decisions, such as which tree species and or size may be selected for harvest. The results of previous studies imply that human management accounts for a greater proportion of change in forest productivity and biomass than does environmental

change (Caspersen et al. 2000; Schimel et al. 2000; Vetter et al. 2005). In interior Alaska, where wildfire is the major source of forest disturbance (Lynch et al. 2002), management decisions directly affect the forest structure through the use of different levels of fire suppression activities across the landscape (Haggstrom 2003).

In interior Alaska there are approximately 9.6 million hectares of commercial forest land and an additional 33.2 million hectares of open woodland that could potentially be harvested or thinned to reduce hazardous fuel conditions (Angelstam and Bergman 2004; Van Cleve et al. 1983). While the United States is one of the chief lumber-producing countries in the world (Houghton 2005), interior Alaska is experiencing a growing portion of its timber harvests being used as fuel. This increased interest in bio-fuels and other forms of bio-energy within Alaska is being driven by a combination of energy independence and high energy prices (Parnell 2012). A number of communities within Alaska are beginning to incorporate wood-fired energy systems utilizing cord wood, chips, and wood pellets. While the scale of the current operations is fairly small (from 10's to 100's of hectares), the increased interest in biomass will potentially impact thousands of hectares of forested lands annually (Forestry 2012). The utilization of forest biomass for bio-fuel production may present an opportunity for forest management and rural economic development; it will have impacts on species, habitats, landscapes, and the society as a whole. The direction and scale of any biomass project will need to take these impacts into account during their development.

The research presented in this dissertation bridges the gap in our understanding of forest diversity within the boreal forest of Alaska. In Chapter 2, I investigate the effects of species and tree size diversity on recruitment within the boreal forest. In this study, I had two objectives: 1) to determine whether species and tree size diversity had significant and consistent effects on recruitment and 2) to characterize the magnitude and spatial patterns of these effects across boreal Alaska. In Chapter 3, I investigate the current tree species and tree size-class diversity within the boreal forest of Alaska and predict possible future scenarios for these two diversity measures. I then develop a spatially dynamic model depicting forest diversity for the Alaskan boreal forest. Lastly, in Chapter



4, I develop a spatial model depicting aboveground forest biomass for the Alaskan boreal forest using a suite of environmental predictors including species and tree size diversity to develop stand-level predictions.

This research will aid forest managers in making more informed decisions in order to maximize sustainable forestry operations in the face of change. The boreal forest and the Alaska residents are faced with many current and future challenges. Climate change and the rising costs of energy are affecting the citizens and the forest in profound ways. A growing interest in the utilization of forest biomass may represent a real opportunity for forest management within this region, which has never experienced industrial-scale forestry. While this may present several new opportunities for hazardous fuel reduction projects and increased employment, it will also present new challenges. The material presented in this dissertation will help shape the future of forestry within the State of Alaska.

## 1.2 REFERENCES

- Alban, D.H., G.E. Host, J.D. Elioﬀ, and D. Shadis. 1994. Soil and vegetation response to soil compaction and forest floor removal after aspen harvesting. US Department of Agriculture, Forest Service, Res. Pap. NC-315, North Central Forest Experiment Station, St. Paul, MN:8.
- Anderson, R., O. Loucks, and A. Swain. 1969. Herbaceous response to canopy cover, light intensity, and throughfall precipitation in coniferous forests. *Ecology* 50(2):255-263.
- Angelstam, P., and P. Bergman. 2004. Assessing Actual Landscapes for the Maintenance of Forest Biodiversity: A Pilot Study Using Forest Management Data. *Ecological Bulletins*(51):413-425.
- Bakker, J.P., H. Olff, J.H. Willems, and M. Zobel. 1996. Why do we need permanent plots in the study of long-term vegetation dynamics? *J. Veg. Sci.* 7(2):147-156.
- Bengston, D.N. 1994. Changing Forest Values and Ecosystem Management. *Soc Natur Resour* 7(6):515-533.
- Bormann, B.T., R.W. Haynes, and J.R. Martin. 2007. Adaptive Management of Forest Ecosystems: Did Some Rubber Hit the Road? *Bioscience* 57(2):186-191.
- Brundtland, G.H. 1987. World commission on environment and development. Our common future. Oxford University Press, Oxford, United Kingdom.
- Buongiorno, J., and J.K. Gilles. 2003. Decision methods for forest resource management. Academic Press, San Diego, CA. 439 p.
- Burton, P.J., C. Messier, D.W. Smith, and W.L. Adamowicz (eds.). 2003. Towards sustainable management of the boreal forest. NRC Research Press, Ottawa, Ontario, Canada. 1039 p.

Caspersen, J.P., S.W. Pacala, J.C. Jenkins, G.C. Hurtt, P.R. Moorcroft, and R.A. Birdsey. 2000. Contributions of land-use history to carbon accumulation in US forests. *Science* 290(5494):1148-1151.

Charron, M. 2005. Sustainable forest management in Canada: Clear policy—Questionable practice. PRB 05-13E Library of Parliament, P.I.a.R.S., Canada. (ed.). Available online at [www.parl.gc.ca/information/library/PRBpubs/prb0513-e.htm](http://www.parl.gc.ca/information/library/PRBpubs/prb0513-e.htm); last accessed May 2008.

Duvall, M.D., and D.F. Grigal. 1999. Effects of timber harvesting on coarse woody debris in red pine forests across the Great Lakes states, USA. *Can. J. For. Res.-Rev. Can. Rech. For.* 29(12):1926-1934.

Eriksson, S., and M. Hammer. 2006. The challenge of combining timber production and biodiversity conservation for long-term ecosystem functioning—A case study of Swedish boreal forestry. *For. Ecol. Manage.* 237(1):208-217.

Fassnacht, K.S., W.B. Cohen, and T.A. Spies. 2006. Key issues in making and using satellite-based maps in ecology: A primer. *For. Ecol. Manage.* 222(1-3):167-181.

Forestry, A.B.o. 2012. Minutes from August 15-16, 2012. P. 14 *in* Alaska Board of Forestry. Alaska Board of Forestry, DOF Conference Room, Palmer, Alaska.

Foster, B.C., D. Wang, W.S. Keeton, and M.S. Ashton. 2010. Implementing sustainable forest management using six concepts in an adaptive management framework. *Journal of Sustainable Forestry* 29(1):79-108.

Fu, S.L., F.W. Bell, and H.Y.H. Chen. 2007. Long-term effects of intensive silvicultural practices on productivity, composition, and structure of northern temperate and boreal plantations in Ontario, Canada. *For. Ecol. Manage.* 241(1-3):115-126.

Haggstrom, D.A. 2003. Managing fire In the urban interface of interior Alaska. P. 16-20 in Proceedings of the 2nd International Wildland Fire Ecology and Fire Management Congress.

Holling, C.S. (e.d.) 1978. Adaptive environmental assessment and management. John Wiley & Sons, New York, New York, USA.

Homeier, J., S.W. Breckle, S. Gunter, R.T. Rollenbeck, and C. Leuschner. 2010. Tree Diversity, Forest Structure and Productivity along Altitudinal and Topographical Gradients in a Species-Rich Ecuadorian Montane Rain Forest. *Biotropica* 42(2):140-148.

Houghton, R.A. 2005. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* 11(6):945-958.

Huettmann, F. 2007. Modern adaptive management: adding digital opportunities towards a sustainable world with new values. *Forum Public Policy* 3:337-342.

Iverson, L.R., and A.M. Prasad. 2001. Potential changes in tree species richness and forest community types following climate change. *Ecosystems* 4(3):186-199.

Lei, X.D., W.F. Wang, and C.H. Peng. 2009. Relationships between stand growth and structural diversity in spruce-dominated forests in New Brunswick, Canada. *Can. J. For. Res.-Rev. Can. Rech. For.* 39(10):1835-1847.

Liang, J.J., J. Buongiorno, R.A. Monserud, E.L. Kruger, and M. Zhou. 2007. Effects of diversity of tree species and size on forest basal area growth, recruitment, and mortality. *For. Ecol. Manage.* 243(1):116-127.

Lin, C.R., J. Buongiorno, J. Prestemon, and K. Skog. 1998. Growth model for uneven-aged loblolly pine stands - simulations and management implications. US Department of Agriculture, Forest Service Res. Pap. FPL-RP-569, Forest Products Lab., Madison, WI:13.

- Lusk, C.H., and A. Ortega. 2003. Vertical structure and basal area development in second-growth *Nothofagus* stands in Chile. *J. Appl. Ecol.* 40(4):639-645.
- Lynch, J.A., J.S. Clark, N.H. Bigelow, M.E. Edwards, and B.P. Finney. 2002. Geographic and temporal variations in fire history in boreal ecosystems of Alaska. *J. Geophys Res-Atmos* 108(D1):8152.
- Malone, T., J. Liang, and E.C. Packee. 2009. Cooperative Alaska Forest Inventory. Gen. Tech. Rep. PNW-GTR-785. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 42 p.
- McRoberts, R.E., S. Winter, G. Chirici, E. Hauk, D.R. Pelz, W.K. Moser, and M.A. Hatfield. 2008. Large-scale spatial patterns of forest structural diversity. *Can. J. For. Res.-Rev. Can. Rech. For.* 38(3):429-438.
- Nadrowski, K., C. Wirth, and M. Scherer-Lorenzen. 2010. Is forest diversity driving ecosystem function and service? *Current Opinion in Environmental Sustainability* 2(1-2):75-79.
- Oren, R., R.H. Waring, S.G. Stafford, and J.W. Barrett. 1987. Twenty-four years of ponderosa pine growth in relation to canopy leaf area and understory competition. *For. Sci.* 33(2):538-547.
- Parnell, S. 2012. Working toward 50% renewables by 2025. Alaska Power Association 61st Annual Meeting.
- Poso, S. 2006. Change monitoring with permanent sample plots. P. 65-84 in *Forest inventory, methodology and applications*, Kangas, A., and M. Maltamo (eds.). Springer, Amsterdam, Netherlands.
- Ruefenacht, B., M.V. Finco, M.D. Nelson, R. Czaplewski, E.H. Helmer, J.A. Blackard, G.R. Holden, A.J. Lister, D. Salajanu, D. Weyermann, and K. Winterberger. 2008.

Conterminous US and Alaska Forest Type Mapping Using Forest Inventory and Analysis Data. *Photogramm. Eng. Remote Sens.* 74(11):1379-1388.

Schimel, D., J. Melillo, H. Tian, A.D. McGuire, D. Kicklighter, T. Kittel, N. Rosenbloom, S. Running, P. Thornton, D. Ojima, W. Parton, R. Kelly, M. Sykes, R. Neilson, and B. Rizzo. 2000. Contribution of increasing CO<sub>2</sub> and climate to carbon storage by ecosystems in the United States. *Science* 287(5460):2004-2006.

Schwartz, J.W., L.M. Nagel, and C.R. Webster. 2005. Effects of uneven-aged management on diameter distribution and species composition of northern hardwoods in Upper Michigan. *For. Ecol. Manage.* 211(3):356-370.

Shao, G.F., H. Wang, L.M. Dai, G. Wu, Y.S. Li, R.L. Lang, and B. Song. 2005. Integrating stand and landscape decisions for multi-purposes of forest harvesting. *For. Ecol. Manage.* 207(1-2):233-243.

Sheppard, S.R.J. 2005. Participatory decision support for sustainable forest management: a framework for planning with local communities at the landscape level in Canada. *Can. J. For. Res.-Rev. Can. Rech. For.* 35(7):1515-1526.

Smith, D.M., B.C. Larson, M.J. Kelty, and P.M.S. Ashton. 1997. The practice of silviculture: applied forest ecology. John Wiley & Sons.

Stott, C.B. 1947. Permanent growth and mortality plots in half the time. *J. For.* 45(9):669-673.

Tilman, D., P.B. Reich, J. Knops, D. Wedin, T. Mielke, and C. Lehman. 2001. Diversity and productivity in a long-term grassland experiment. *Science* 294(5543):843-845.

Van Cleve, K., C.T. Dyrness, L.A. Viereck, J. Fox, F.S. Chapin, and W. Oechel. 1983. Taiga ecosystems in interior Alaska. *Bioscience* 33(1):39-44.

Varga, P., H.Y.H. Chen, and K. Klinka. 2005. Tree-size diversity between single- and mixed-species stands in three forest types in western Canada. *Can. J. For. Res.-Rev. Can. Rech. For.* 35(3):593-601.

Vetter, M., C. Wirth, H. Bottcher, G. Churkina, E.D. Schulze, T. Wutzler, and G. Weber. 2005. Partitioning direct and indirect human-induced effects on carbon sequestration of managed coniferous forests using model simulations and forest inventories. *Glob. Change Biol.* 11(5):810-827.

Viereck, L., and E. Little. 2007. *Alaska trees and shrubs*. University of Alaska Press, Fairbanks, Alaska, USA.

Vila, M., J. Vayreda, L. Comas, J.J. Ibanez, T. Mata, and B. Obon. 2007. Species richness and wood production: a positive association in Mediterranean forests. *Ecol Lett* 10(3):241-250.

Vila, M., J. Vayreda, C. Gracia, and J.J. Ibanez. 2003. Does tree diversity increase wood production in pine forests? *Oecologia* 135(2):299-303.

Von Gadow, K., M. Tomé, and T. Pukkala (eds.). 2001. *Sustainable Forest Management*. Kluwer Academic Print on Demand.

Wagner, S. 2004. Möglichkeiten und Beschränkungen eines funktionsorientierten Waldbaus. *Forst und Holz* 59:105-111.

Walters, C. 1986. *Adaptive management of renewable resource*. Macmillan Publishing, New York, New York, USA.

Walters, C.J., and C.S. Holling. 1990. Large-Scale Management Experiments and Learning by Doing. *Ecology* 71(6):2060-2068.

**Wang, S. 2004. One hundred faces of sustainable forest management. *Forest Policy and Economics* 6(3-4):205-213.**

**Wollenberg, E., D. Edmunds, and L. Buck. 2000. *Anticipating Change: Scenarios as a tool for adaptive forest management: A guide*. Center for International Forestry Research (CIFOR), Bogor, Indonesia.**



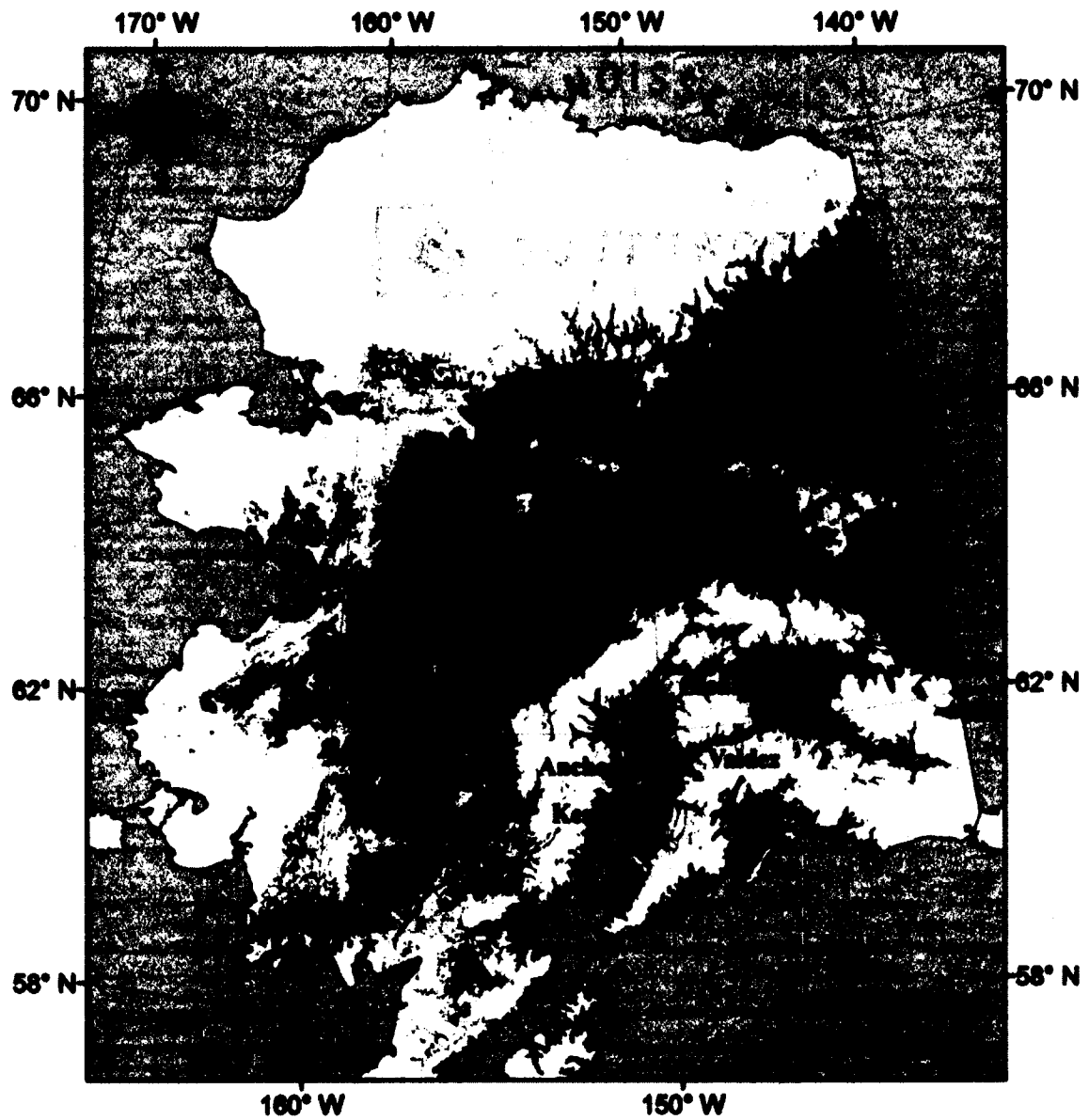


Figure 1.1: Geographic distribution of the 704 Sample Plots (in triangles) within the Alaskan boreal forest (Ruefenacht et al., 2008).

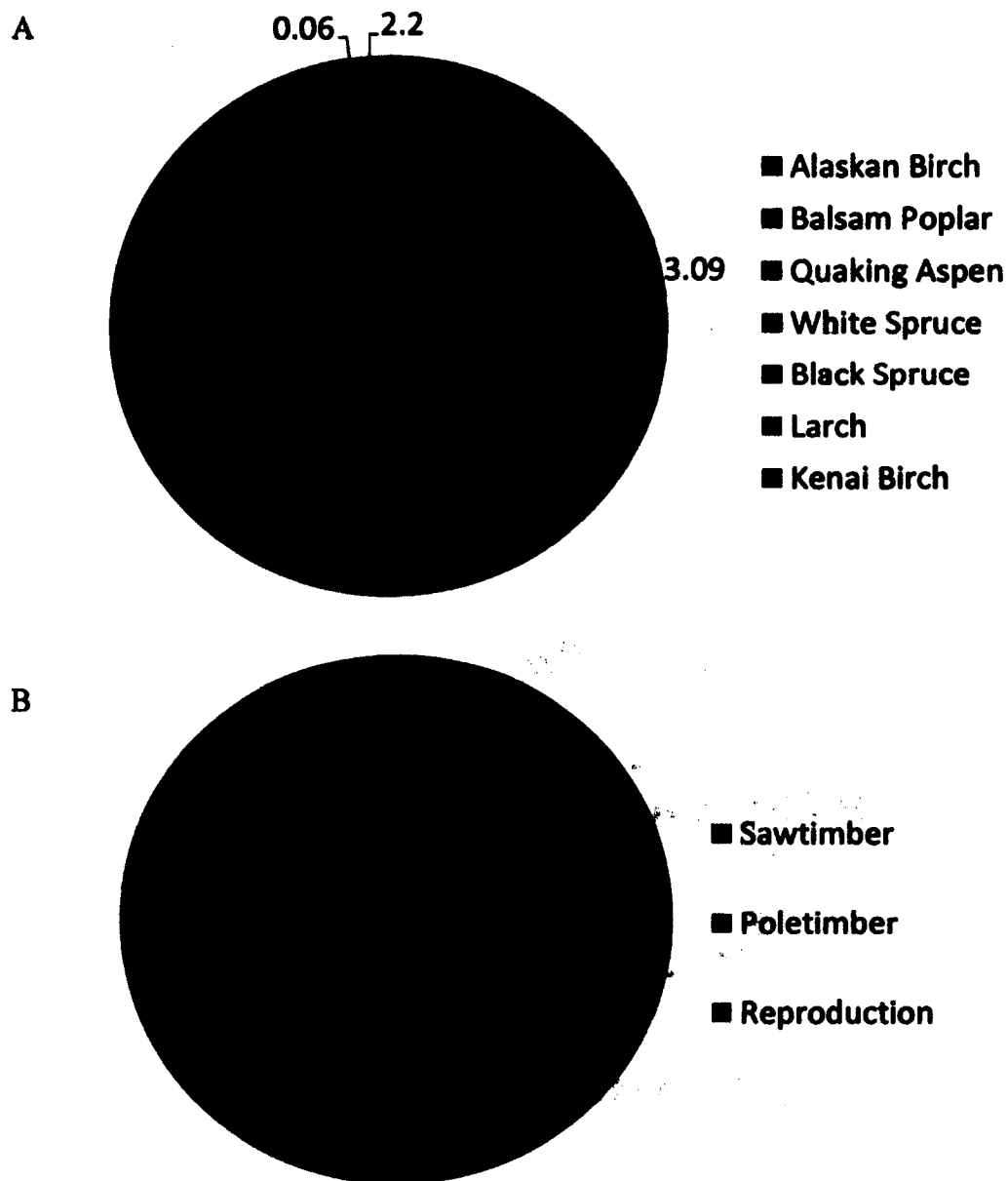


Figure 1.2: Distribution of species by volume of total growing stock (A) and stand diameter breast height (dbh) size classes (B) in the forests of interior Alaska. Sawtimber: 22.9cm DBH or greater, Poletimber: 15cm to 22.8cm DBH, Reproduction: 2.5cm to 14.9cm DBH (Malone et al., 2009).

## CHAPTER 2: EFFECTS OF SPECIES AND TREE SIZE DIVERSITY ON RECRUITMENT IN THE ALASKAN BOREAL FOREST: A GEOSPATIAL APPROACH<sup>2.1</sup>

### 2.1 ABSTRACT

This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's Index, within mixed poplar/birch (*Populus tremuloides*, *P. balsamifera* and *Betula neoalaska*) and mixed spruce (*Picea glauca* and *P. marianana*) stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

### 2.2 INTRODUCTION

Understanding the effects of biodiversity on forest recruitment poses a challenge because recruitment is subject to both non-spatial and spatial effects of various biotic and

---

<sup>2.1</sup> Published in slightly modified form as: Young B, Liang JJ, Chapin FS. Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach. *For. Ecol. Manage.* (2011) 262:1608-1617.

abiotic factors, such as stand density, climate, and soil fertility (Caspersen and Pacala, 2001; Mladenoff, 2005; Pretzsch, 2005; Turner, 2005; Liang *et al.*, 2007). Previous studies, which have only addressed the non-spatial effects, have found a positive relationship between species diversity and recruitment in a variety of forest types (Liang *et al.*, 2007; Lei *et al.*, 2009). Structural diversity, notably tree size diversity, is another key component of forest biodiversity. Tree size diversity is often manipulated by forest managers to increase biodiversity (Buongiorno *et al.*, 1994; Buongiorno and Gilles, 2003; McRoberts *et al.*, 2008). The effect of tree size diversity on recruitment, like that of species diversity, is still rather poorly understood (Liang *et al.*, 2007; Lei *et al.*, 2009). Positive, negative, and insignificant effects on recruitment have been reported (Don *et al.*, 2007; Liang *et al.*, 2007; Lei *et al.*, 2009), although none of these previous studies took spatial autocorrelation into account.

Biodiversity goals are commonly used as the basis for management decisions. Different measures of biodiversity, however, may support different solutions. Several indices of biodiversity have been applied in previous studies evaluating the effects of biodiversity on forest productivity (Lindenmayer *et al.*, 2000; Liang *et al.*, 2007; McRoberts *et al.*, 2008; Lei *et al.*, 2009). Two of the most common indices used in forest biodiversity studies are the Shannon's index (Shannon, 1948) and the Simpson's index (Simpson, 1949). Although these two indices both depend on landscape richness and evenness, they weight rare classes differently (Magurran, 2004). The Shannon's index reflects both evenness and richness of classes by weighing all classes in proportion to their frequencies in the sample (Magurran, 1988; Jost, 2006) however, it has been shown to be overly sensitive to rare classes in the population (Magurran, 2004). In contrast, the Simpson's index is heavily weighted towards the most abundant classes in the sample, it is less sensitive to the number of different classes (Magurran, 2004). These differences between the two indices may result in different relationships between diversity and forest recruitment.

The spatial effects of biodiversity on recruitment have rarely been addressed in previous studies of recruitment and biodiversity, especially at landscape and regional

scales (Roberts and Gilliam, 1995; Bond and Chase, 2002; West *et al.*, 2009). To this end, spatial autocorrelation, a general property of most ecological attributes due to physical or community processes (Legendre, 1993; Bivand *et al.*, 2008), is a key issue to address, especially in large-scale forest studies (Liang and Zhou, 2010). When unaccounted for, spatial autocorrelation may affect statistical model predictions because it violates the assumption of independence on which most standard statistical procedures rely (Legendre, 1993). Geospatial models that account for spatial autocorrelation could be useful in assessing the spatial effects of biodiversity on recruitment, especially when specific environmental drivers such as temperature and precipitation are not included in the model (Bivand *et al.*, 2008).

Two general hypotheses — sampling effect and niche complementarity — have been proposed to explain the relationships between biodiversity and forest recruitment. The sampling effect hypothesis suggests that biodiversity increases ecosystem productivity through one or a few dominant, high-biomass species in the polyculture (Tilman *et al.*, 1997; Cardinale *et al.*, 2006; Fargione *et al.*, 2007). The niche complementarity hypothesis states that biodiversity enhances ecosystem productivity because niche differences among species and tree size groups enable the forest community to access larger quantities of limiting resources (Loreau and Hector, 2001; Tilman *et al.*, 2001).

Two phases of recruitment are particularly important to stand development in the boreal forest: (1) seedling recruitment and (2) growth of seedlings and sprouts to the sapling stage (Zasada, 1986). Seedling recruitment is controlled mainly by seedbed properties (e.g., thickness, temperature, and moisture) rather than by composition of the biotic community. The second phase of recruitment, where seedlings and sprouts grow into the sapling stage, involve competitive interactions that determine which species dominate the overstory after a disturbance (Greene *et al.*, 1999). The sapling phase of recruitment depends on both biotic and abiotic properties of the forest stand (Kneeshaw and Bergeron, 1998; Greene *et al.*, 1999; Messier *et al.*, 1999; McCarthy, 2001). This article focuses on this second phase of tree recruitment.

The Alaskan boreal forest extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains (Figure 2.1), covering an area of nearly 500,000 km<sup>2</sup>. The boreal forest in Alaska consists of a mosaic of two general forest types, mixed poplar/birch and mixed spruce (Viereck and Little, 2007; Ruefenacht *et al.*, 2008) and has only eight tree species. White spruce (*Picea glauca*) and black spruce (*P. marianana*) are the predominant conifers and two poplars (*Populus tremuloides* and *P. balsamifera*) and Alaskan birch (*Betula neoalaska*) represent the majority of the deciduous trees. The floristic simplicity of the boreal forest within Alaska makes it much easier to study the effects of biodiversity on tree recruitment.

The main goal of this paper is to investigate the effects of species and tree size diversity on recruitment within the boreal forest. We have two objectives: to determine if species and tree size diversity had significant and consistent effects on recruitment for the two general forest types that occur in Alaska and to characterize the magnitude and spatial patterns of these effects across boreal Alaska.

## 2.3 DATA AND METHODS

### 2.3.1 DATA

The data were obtained from the Cooperative Alaska Forest Inventory (CAFI) Database (Malone *et al.*, 2009), which consists of field-gathered information from over 600 periodically re-measured permanent sample plots across interior and south-central Alaska north of 60°N (Figure 2.1). Most permanent sample plots occur on sites that are potentially suitable for commercial harvest and are well stocked and road-accessible. All plots can be categorized as one of the two forest group types recognized by the USDA Forest Service forest inventory and analysis program for the Alaskan boreal forest: Aspen/Birch and Spruce/Fir (Figure 2.1; Ruefenacht *et al.*, 2008). In our study, we assigned each permanent sample plot to a forest group type (Aspen/Birch or Spruce/Fir) based on the number of trees of each species in the plot (Ruefenacht *et al.* 2008). In our

study, we refer to the Aspen/Birch group as mixed poplar/birch because some plots were dominated by *P. balsamifera*. In addition, we refer to the Spruce/Fir group in our study as mixed spruce due to the lack of fir in our study region.

All the CAFE plots are square in shape and 0.04 ha (0.10 ac) in size. Most of them have been measured at least twice at 5-year intervals since 1994. In order to minimize the effect of human and natural disturbances such as logging, wildfire, and insect damage, we used data only from those plots with less than 50 percent tree mortality between two consecutive inventories. In total, 438 plots were selected of which 212 were classified as mixed poplar/birch and 226 as mixed spruce (Table 2.1). For each plot, we calculated the mean annual recruitment,  $r$ , expressed in basal area (Table 2.2). Stand recruitment basal area was calculated by summing the basal area of those trees that outgrew the 2.54 cm diameter threshold between two consecutive inventories.

Two different diversity indices, Shannon's index (Shannon, 1948) and the Simpson index (Simpson, 1949), were used in this study to assess tree size and species diversity. The Shannon's index was used to calculate tree species diversity,  $H_s$ , and tree size diversity,  $H_d$  on each plot:

$$H_s = - \sum_{i=1}^{ns} \frac{B_i}{B} \cdot \ln \left( \frac{B_i}{B} \right) \quad (2.1)$$

and

$$H_d = - \sum_{j=1}^{nd} \frac{B_j}{B} \cdot \ln \left( \frac{B_j}{B} \right) \quad (2.2)$$

and the inverse of the Simpson index (Simpson, 1949) was used so that increasing index values were associated with increasing diversity to calculate tree species diversity,  $D_s$ , and tree size diversity,  $D_d$  on each plot:

$$D_s = 1 - \sum_{i=1}^{ns} \left( \frac{B_i}{B} \right)^2 \quad (2.3)$$

and

$$D_d = 1 - \sum_{j=1}^{nd} \left( \frac{B_j}{B} \right)^2 \quad (2.4)$$

for both indices  $B$ ,  $B_i$ , and  $B_j$  were, respectively, the total stand basal area, the basal area of trees of species  $i$ , and the basal area of trees in diameter class  $j$ ; and  $ns$  and  $nd$  represented total number of tree species and diameter classes. With 7 species (Table 1) and 20 2.54 cm diameter classes being studied here,  $ns = 7$  and the theoretical range of  $H_s$  was between 0 and  $\ln(7) = 1.95$  while  $D_s$  was between 0 and 1, whereas  $nd = 20$  and the theoretical range of  $H_d$  was between 0 and  $\ln(20) = 3.00$  and  $D_d$  was between 0 and 1. There was little correlation between  $H_s$  and  $H_d$  ( $P = 0.23$ ) or  $D_s$  and  $D_d$  ( $P = 0.35$ ); thus species diversity and tree size diversity represented two distinct measures of forest diversity when the Shannon diversity index was used.

We tested Simpson's index of tree size and species diversity but found that it was very strongly correlated with other variables in the analysis and therefore did not provide an assessment of diversity that was independent of other potential explanatory variables. It was therefore less useful than Shannon's index in examining the independent effects of tree size and species diversity on recruitment patterns. This is most likely a result of Shannon's index weighing all classes in proportion to their frequencies in the sample (Magurran, 1988; Jost, 2006) while Simpson's index heavily weights towards the most abundant classes in the sample (Magurran, 2004). For this reason, we present only the results of Shannon's index.

In addition to the diversity measures, we studied 11 other spatially explicit variables that might influence recruitment. Stand basal area,  $B$ , negatively affects recruitment in many forest types (e.g. Lusk and Ortega 2003, Liang *et al.* 2005, Yang *et al.* 2009). Stand age,  $A$ , which was determined from the average age of five of the largest representative individuals within the stand (Tom Malone, University of Alaska Fairbanks, personal communication), was used to represent forest successional status (Caspersen, 2004). Mean diameter at breast height (dbh) of live trees,  $D$ , another versatile measure of forest successional status especially for uneven-aged stands, has been documented to directly influence growth and productivity of many different forest types (e.g. Gower *et*



*al.* 1996, Harper *et al.* 2006, Liang *et al.* 2007). Site index has not been measured on CAFI plots due to a lack of site index models for the region. We therefore created a unitless measure of site productivity,  $P$ , calculated using a function of stand elevation,  $z$ , aspect,  $\alpha$ , and slope,  $s$  (Stage and Salas 2007, p.487):

$$P(s, \alpha, z) = s[c_1 + c_2 \cos(\alpha) + c_3 \sin(\alpha)] + \ln(z+1) \cdot s[c_4 + c_5 \cos(\alpha) + c_6 \sin(\alpha)] + z^2 \cdot s[c_7 + c_8 \cos(\alpha) + c_9 \sin(\alpha)] + c_{10}z + c_{11}z^2 \quad (2.5)$$

where  $c$ 's were parameters to be estimated in each equation.

Forest recruitment is also subject to the influence of other site characteristics. Charcoal,  $C$ , was studied here due to its potential benefits on plant productivity (Wardle *et al.*, 1998; Naydenov *et al.*, 2006). Coarse woody debris (CWD) can exhibit negative effects on seedling development when it is over-abundant (Chen and Popadiouk, 2002), but it has also been shown to increase forest nutrient recycling and affect plant establishment and growth as well as providing important habitat for a variety of insects, invertebrates, birds, fungi, and epiphytes (Franklin, 1981; Molofsky and Augspurger, 1992; Brassard and Chen, 2006, 2008). CWD in our study was represented by standing deadwood or snags,  $G$ , large-diameter ( $> 2.5$  cm) downed woody debris,  $W$ , small-diameter ( $< 2.5$  cm) woody debris, and leaf litter,  $L$ . Additionally, snags,  $G$ , represent an opening of the canopy that can then be occupied by new recruits (Kneeshaw and Bergeron, 1998; Harper *et al.*, 2006). We also documented the total depth of organic matter,  $O$ , as it is generally found to affect germination rates of a variety of boreal forest tree species (Zasada *et al.*, 1992; Johnstone and Chapin, 2006). In addition, easting ( $\lambda$ ) and northing ( $\phi$ ) coordinates of the Universal Transverse Mercator system (UTM) were obtained for the center of each permanent sample plot to facilitate calculation of distances among plots.

### 2.3.2 METHODS

Two phases of recruitment are particularly important to stand development in the boreal forest: seedling recruitment and growth of seedlings and sprouts to the sapling stage (Zasada, 1986). Seedling recruitment is mainly controlled by seedbed properties (e.g., thickness, temperature, and moisture) rather than composition of the biotic community. In the second phase of recruitment, seedlings and sprouts grow into the sapling stage and have the potential to replace the overstory following a disturbance (Greene *et al.*, 1999). This phase of recruitment depends on both biotic and abiotic properties of the forest stand (e.g. (Kneeshaw and Bergeron, 1998)). This article focuses on this second phase of tree recruitment.

The factors that most strongly influenced recruitment were selected for the two forest types using the hierarchical partitioning (HP) method (Chevan and Sutherland, 1991) to overcome erroneous estimates of levels of significance caused by multicollinearity (Mac Nally, 1996). The HP method helps to identify the most influential predictor variables by capturing their independent and joint contribution to the goodness-of-fit of recruitment (Chevan and Sutherland, 1991; Mac Nally, 2000). In our analysis, the coefficient of determination,  $R^2$ , was used as the goodness-of-fit measure, and the HP procedures were conducted using the hier.part package ver.1.0 (Walsh and Mac Nally, 2009) of the R system (Team, 2010). Statistical significance of each variable was then approximated with a randomization procedure based on the Z-score with an upper 95% confidence limit ( $[Z] \leq 1.65$ ) (Mac Nally, 2002).

Having identified the most influential factors of recruitment for each forest type using HP, we estimated linear recruitment models with the Generalized Least-Squares (GLS) method:

$$\mathbf{r} = \boldsymbol{\beta} \cdot \mathbf{X} + \mathbf{e}(\lambda, \varphi) \quad (2.6)$$

where  $\mathbf{r}$  was the vector of annual recruitment ( $\text{m}^2\text{ha}^{-1}\text{year}^{-1}$ ).  $\mathbf{X}$  and  $\boldsymbol{\beta}$  represented the matrix of explanatory variables selected by the HP procedures, and the vector of

coefficients estimated with GLS, respectively. The normally distributed random residuals  $e(\lambda, \phi)$ , if shown to be spatially autocorrelated, were assumed to observe an isotropic and spherical empirical semivariogram (Cressie, 1993).

We tested for spatial autocorrelation as well as for large-scale spatial patterns of residuals  $e_k(\lambda, \phi)$ , assuming that plots at distant locations will affect each other less than at nearby locations, using a spatial weight of inverse distance. Given the neighborhood structure, we then evaluated the residuals of the linear recruitment models using Moran's  $I$  and Geary's  $C$  test statistics (Sokal and Oden 1978) with the *spdep* (Bivand *et al.*, 2007) package for the R system.

If spatial autocorrelation was found to be significant in the residuals of the linear recruitment models, we then assessed the effects of species and tree size diversity across the landscape on recruitment with the following universal kriging model (Cressie, 1993; Ver Hoef, 1993):

$$\mathbf{r} = \boldsymbol{\alpha} \cdot \mathbf{X} + \boldsymbol{\delta} \quad (2.7)$$

where  $\boldsymbol{\alpha}$  was a vector of parameters to be estimated for the matrix of selected variables  $\mathbf{X}$ , and their product represented a non-spatial trend.  $\boldsymbol{\delta}$  was a zero-mean intrinsically stationary random process with a semivariogram estimated with restricted maximum likelihood (REML; Bartlett, 1937). This universal kriging model simultaneously fit the non-spatial trends in the data through trend surface regression and variogram fitting (Ver Hoef, 1993).

In order to conduct the universal kriging models, we first created a grid of points with a 1 - km<sup>2</sup> resolution for each of the two forest types (Ruefenacht *et al.*, 2008) represented in this study using ArcGIS 10.0 (ESRI, Redlands, CA). For each grid point, the mean values for each of the different trend variables were then assigned. The trend variables used in the universal kriging model were previously determined by the HP method. The semivariogram and kriging models were fit with the *gstat* package (Pebesma 2004) for the R system. The predicted values generated from the kriging models were then mapped using ArcGIS 10.0.

To analyze the sensitivity of recruitment to changes in species and tree size diversity, we generated three different recruitment scenarios by using high (mean + 2 × standard deviation), mean, and low (mean – 2 × standard deviation) values for both species and tree size diversity, while holding all other factors fixed at their sample mean (Table 2.2). The sensitivity analysis was conducted to investigate the biological importance of species and tree size diversity across the landscape in addition to their statistical significance as identified by the REML estimation (Bartlett, 1937).

## 2.4 RESULTS

### 2.4.1 ATTRIBUTES OF THE TWO FOREST TYPES

The two forest types, mixed poplar/birch and mixed spruce, display a species frequency distribution that differs strikingly between all live trees and recruits (Table 2.1). For the mixed poplar/birch forest type, 59% of all live trees are deciduous, whereas only 4% of the recruits are deciduous. In the mixed spruce forest type, 42% of all live trees are represented by deciduous species while 9% of the recruits are deciduous. *P. glauca* represents over three quarters of the recruits in mixed poplar/birch stands and over half in the mixed spruce stands. Among conifers, *P. mariana* is twice as frequent among recruits in mixed spruce compared to mixed poplar/birch stands. The quantity of deciduous recruits is minimal with the noted exception of *B. kenaica*, which accounts for just over 5% of the recruits in the mixed spruce stands. *B. kenaica* accounts for 63% of recruitment in stands where it is found on the Kenai Peninsula (see Figure 2.1) and is absent from most other stands.

The mixed poplar/birch and the mixed spruce forest types were relatively similar to one another in most attributes (Table 2.2). Compared to mixed spruce stands, however, mixed poplar/birch stands had significantly greater (*t*-tests with unequal variances) site productivity, *P* ( $P < 0.001$ ), depth of organic matter, *O* ( $P = 0.038$ ), easting of UTM coordinates,  $\lambda$  ( $P < 0.001$ ), and charcoal, *C* ( $P = 0.016$ ). However, mixed poplar/birch

stands had significantly lower deadwood,  $L$  ( $P < 0.001$ ), and average stand age,  $A$  ( $P = 0.028$ ). None of the other attributes differed significantly ( $p > 0.05$ ) between forest types.

#### 2.4.2 FACTOR IDENTIFICATION AND PARAMETER ESTIMATION

The factors that most strongly influenced recruitment were identified for both forest types using the HP method (Table 2.3). For the mixed poplar/birch forest type, tree size diversity ( $H_d$ ) and species diversity ( $H_s$ ) both had significant effects on recruitment and contributed 12.3% and 5.0 %, respectively, toward the goodness of fit. In addition, stand age ( $A$ ), stand mean diameter ( $D$ ), the amount of charcoal in the soil ( $C$ ), and the number of snags ( $G$ ) significantly contributed for a combined contribution of 68.9% toward the goodness of fit (Table 2.3).

For the mixed spruce forest type, there was an independent contribution of 12.9% for tree size diversity ( $H_d$ ) and 7.9% for species diversity ( $H_s$ ) (Table 3). In addition, two other stand-level variables, stand basal area ( $B$ ) and stand mean diameter ( $D$ ), significantly contributed 17.9% and 41.2% respectively toward a combined 59.1% goodness of fit for recruitment (Table 2.3).

Recruitment was then estimated with parsimonious linear models that contained the significant explanatory variables identified by the HP method (Models (a) in Table 2.4). Some of the variables with high contribution to the goodness of fit were found to be non-significant in the linear models (Table 2.4), perhaps due to the correlation of tree size and species diversity with other predictor variables (Mac Nally, 2002), and a significant spatial autocorrelation in the residuals of the parsimonious models for both mixed poplar/birch ( $P < 0.001$ ) and mixed spruce stands ( $P < 0.001$ ) (Table 2.6).

When spatial autocorrelation had been accounted for, the geospatial recruitment models (Eq. 2.5; Models (b) in Table 2.4) differed significantly from the non-spatial models (Eq. 2.4 without the error term; Models (a) in Table 2.4) in terms of coefficients and goodness-of-fit. The geospatial models were of better quality as judged by the more negative AIC and BIC values (Table 2.4). We therefore chose the geospatial models

(Models (b) in Table 2.4) to study the effects of species and tree size diversity on recruitment for both forest types in Alaska.

### 2.4.3 EFFECTS OF TREE SIZE AND SPECIES DIVERSITY ON RECRUITMENT

#### 2.4.3.1 TREE SIZE DIVERSITY

The predicted effect of tree size diversity on recruitment for the mixed poplar/birch forest varied greatly across Alaska (Figure 2.2A, B, and C). For high levels of tree size diversity (Figure 2.2A), which negatively affects recruitment (Table 2.4B), the predicted rate of recruitment over a 10-year period for the majority of the region was below  $0.1 \text{ m}^2\text{ha}^{-1}$  with an exception of a few localized areas in Matanuska-Susitna Valley north of Anchorage and the Copper River Valley south of Glennallen. With decreasing levels of tree size diversity (Figure 2.2B, C), the overall levels of recruitment increased across the state with the greatest expected gains occurring in Matanuska-Susitna, Copper River Valleys, and the Tanana Flats area from Tok to Fairbanks. A small pocket of low recruitment persisted on the Kenai Peninsula, regardless of changes in tree size diversity.

Within the mixed spruce forest type, the 10-year recruitment predicted from tree size diversity was higher and more uniform across the study region than with the mixed poplar/birch forest (Figure 2.2D, E, and F) because tree size diversity did not strongly affect recruitment (Table 2.4B). For high levels of tree size diversity (Figure 2.2D), the recruitment for the vast majority of the region was below  $0.3 \text{ m}^2\text{ha}^{-1}$  with only a few small localized areas, such as the Copper River Valley south of Glennallen that experienced higher recruitment. With decreasing levels of tree size diversity (Figure 2.2E and F), the overall levels of recruitment increased and became nearly homogeneous across the state with the greatest expected recruitment occurring on a corridor from the Copper River Valley to Anchorage.

The distribution of the average basal area of trees across the 20 different size classes differs between stands with high ( $H_d > 2.3$ ) and low ( $H_d < 1.3$ ) tree size diversity within both forest types (Figure 2.3A and B). Within the mixed poplar/birch forest type

(Figure 3a), the low-diversity plots only have trees smaller than size class 11, with the majority being less than size class 5. On the high diversity plots within this forest type, the trees on average are distributed across all the size classes with the majority of the basal area falling in the two largest size classes. For the mixed spruce forest type (Figure 2.3B), within the low diversity plots, the total basal area comes from trees of size class 7 and smaller while, in the high diversity plots, the total basal area is more evenly distributed across the size classes with the largest size class contributing just over 20%. Thus more structurally diverse forest stands produce larger trees and greater stand biomass, at least in terms of tree size classes.

#### 2.4.3.2 TREE SPECIES DIVERSITY

The predicted effect of species diversity on recruitment for the mixed poplar/birch forest varied widely across Alaska (Figure 2.4A, B, and C). Within this forest type, for high levels of species diversity (Fig 2.4A), which tends to promote recruitment (Table 2.4B), the predicted recruitment rates differed greatly across the landscape, with the lowest predicted values ( $0.0 \text{ m}^2 \text{ ha}^{-1}$ ) on the Kenai Peninsula and the highest ( $>0.6 \text{ m}^2 \text{ ha}^{-1}$ ) in the Matanuska-Susitna and Copper River Valleys. With decreasing tree species diversity (Fig 2.3B, C), the overall rate of recruitment declined with areas of high recruitment shrinking in size, and areas of zero recruitment emerging near Fairbanks.

Within the mixed spruce forest type (Figure 2.4D, E, and F); the predicted effect of species diversity on recruitment was far more homogeneous across the landscape. For high levels of species diversity (Figure 2.4D), recruitment for the vast majority of the region was greater than  $0.3 \text{ m}^2 \text{ ha}^{-1}$  with an exception of some small localized regions surrounding Fairbanks, the Matanuska-Susitna Valley, the northern portion of the Kenai Peninsula, and south of Tok. Within the Copper River Valley, high levels ( $>0.6 \text{ m}^2 \text{ ha}^{-1}$ ) of recruitment are predicted regardless of the level of tree species diversity (Figure 2.4D, E, and F). With decreasing levels of species diversity (Figure 2.4E and F), the overall levels of recruitment decreased uniformly across the state with the greatest expected loss occurring on the Kenai Peninsula.

## 2.5 DISCUSSION AND CONCLUSION

This paper presents spatially explicit models that describe the relationships between recruitment and tree size and species diversity for the Alaskan boreal forest. Recruitment was negatively associated with tree size diversity for both the mixed poplar/birch and mixed spruce forest types (Table 2.4). This negative relationship was most likely caused by the reduction in regeneration niches (Grubb, 1977; Denslow, 1987) in many mature boreal forest stands with high tree size diversity (Zasada, 1986; Johnstone *et al.*, 2004) and by the competition for light, water, and other factors (Man *et al.*, 2008), which affect the survival and growth of small seedlings and saplings. Within boreal forest stands under natural succession in the absence of a major disturbance, recruitment is most likely controlled by seedling dynamics, i.e. the diameter growth and mortality (Messier *et al.*, 1999), that contribute most to the sapling recruitment in this study.

Succession in the boreal forest typically leads to a change from deciduous dominance to conifer dominance due to the recruitment of conifers into the canopy (Table 2.1) (Bergeron *et al.*, 2002; Chapin *et al.*, 2006c; Harper *et al.*, 2006). The seedlings that are most likely to grow into the sapling stage should occupy microsites with more favorable light and nutrient regimes. The positive relationship observed in this study between recruitment and the number of snags (Table 2.4) further supports this assumption given that the numbers of snags serve as a proxy for canopy gaps. Post-disturbance recruitment in the boreal forest depends on both the severity and extent of disturbance (Johnstone and Chapin, 2006). Wildfire severity directly affects stand structure and composition because post-fire recruits generally dominate the canopy (Gutsell and Johnson, 2002) for years following the disturbance. However, sites that have experienced high-severity wildfire may also exhibit increased abundance and richness of understory vegetation (Bernhardt *et al.*, 2011), which can negatively affect recruitment (Cater and Chapin, 2000).



While wildfire is the predominant form of disturbance in the boreal forest (Chapin *et al.*, 2006c), it is by no means the only one. For example, between the years 1987 and 2003 the spruce beetle (*Dendroctonus rufipennis*) infested south central Alaska (U.S. Forest Service, 2004) causing massive mortality of smaller-diameter spruce, resulting in stand replacement (Boucher and Mead, 2006). This disturbance increased the abundance and cover of the understory grass species bluejoint (*Calamagrostis canadensis*) (Boucher and Mead, 2006; Boggs *et al.*, 2008). Bluejoint has previously been shown to negatively influence tree recruitment in the boreal forest (Lieffers and Stadt, 1994; Cater and Chapin, 2000). Our findings of lower levels of recruitment on part of the Kenai Peninsula and other locations within Southcentral Alaska matched the patterns of spruce beetle outbreak in these areas.

In contrast to tree size diversity, species diversity had a positive effect on recruitment (Table 2.4), although this effect was not as pronounced as that of tree size diversity (Figure 2.4). These results are consistent with the hypothesis that differences in species life histories lead to complementarity and niche differentiation (Tilman *et al.*, 1997; Loreau *et al.*, 2001; Tilman *et al.*, 2001), resulting in an increase in regeneration niches in species-rich forests (Grubb, 1977). These findings are consistent with previous findings by Caspersen and Pacala (2001), and Liang *et al.*, (2005; 2007) but contrast with the findings by Lei *et al.*, (2009) who found no effects of species diversity on recruitment in the boreal forest of Eastern Canada.

Due to significant spatial variation, the positive effect of species diversity on recruitment was not observed for all locations within the study region. For example, in the eastern interior of Alaska, which is relatively dry (Shulski and Wendler, 2007), recruitment levels were not enhanced with greater levels of species diversity. This may be due to potential stressors of recruitment such as drought (Barrett *et al.*, 2010; Johnstone *et al.*, 2010) despite the facilitative effects of species diversity (Mulder *et al.*, 2001). The breakdown of this facilitative interaction is likely due to the magnitude and source of the particular stressor (Gomez-Aparicio *et al.*, 2008). In addition, these regions may have

historically experienced higher levels of disturbance, which may act synergistically with other stressors to reduce both species diversity and recruitment (Folke *et al.*, 2004).

Results from this study show that tree size diversity and species diversity explained a substantial proportion of the variation in recruitment. The effects of these diversity measures on recruitment are highly variable across the landscape and are thus best modeled using spatial techniques. These variations in the effects were most likely due to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. Our results demonstrated that the positive effect of species diversity, which was also found by Troumbis and Memtsas (2000), Erskine *et al.*, (2006), Liang *et al.*, (2007), and Vila *et al.*, (2007); and the negative effect of tree size diversity, also found by Liang *et al.*, (2007) on recruitment still held when spatial autocorrelation and large-scale spatial patterns had been accounted for.

Considering that both recruitment and biodiversity are spatially driven processes (Roberts and Gilliam, 1995), accounting for spatial autocorrelation is crucial when conducting landscape-level analysis (Sokal and Oden, 1978; Wagner and Fortin, 2005; Bivand *et al.*, 2008). By not taking into account spatial autocorrelation, previous recruitment or biodiversity studies may have generated erroneous results (Legendre, 1993). The method proposed in this study should act as a model to assist forest managers and researchers who are interested in these dynamics.

Forest managers sometimes consider tree size diversity and species diversity as desirable ecological objectives (Buongiorno *et al.*, 1994; Deal, 2007). Our findings suggest that there may be tradeoffs between the two types of diversity that forest managers should consider in promoting forest diversity. Additionally, there are many additional organisms within a forest besides living trees that may influence forest diversity (Franklin *et al.*, 2002). Therefore, further studies should incorporate diversity of the entire plant community, both living and dead, when investigating the relationship between forest productivity and biodiversity.

## 2.6 REFERENCES

- Barrett, K., E.S. Kasischke, A.D. McGuire, M.R. Turetsky, and E.S. Kane. 2010. Modeling fire severity in black spruce stands in the Alaskan boreal forest using spectral and non-spectral geospatial data. *Remote Sens. Environ.* 114(7):1494-1503.
- Bartlett, M.S. 1937. Properties of Sufficiency and Statistical Tests. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences* 160(901):268-282.
- Bergeron, Y., A. Leduc, B.D. Harvey, and S. Gauthier. 2002. Natural fire regime: A guide for sustainable management of the Canadian boreal forest. *Silva Fennica* 36(1):81-95.
- Bernhardt, E.L., T.N. Hollingsworth, and F.S. Chapin. 2011. Fire severity mediates climate-driven shifts in understorey community composition of black spruce stands of interior Alaska. *J. Veg. Sci.* 22(1):32-44.
- Bivand, R.S., L. Anselin, O. Berke, A. Bernat, M. Carvalho, Y. Chun, C.F. Dormann, S. Dray, R. Halbersma, and N. Lewin-Koh. 2007. *spdep: Spatial dependence: weighting schemes, statistics and models*. R package version 0.4-9.
- Bivand, R.S., E.J. Pebesma, and V. Gómez-Rubio. 2008. *Applied spatial data analysis with R*. Springer Verlag, New York, New York, USA.
- Boggs, K., M. Sturdy, D.J. Rinella, and M.J. Rinella. 2008. White spruce regeneration following a major spruce beetle outbreak in forests on the Kenai Peninsula, Alaska. *For. Ecol. Manage.* 255(10):3571-3579.
- Bond, E., and J. Chase. 2002. Biodiversity and ecosystem functioning at local and regional spatial scales. *Ecol. Lett.* 5(4):467-470.

Boucher, T.V., and B.R. Mead. 2006. Vegetation change and forest regeneration on the Kenai Peninsula, Alaska following a spruce beetle outbreak, 1987–2000. *For. Ecol. Manage.* 227(3):233-246.

Brassard, B.W., and H.Y.H. Chen. 2006. Stand structural dynamics of North American boreal forests. *Crit. Rev. Plant Sci.* 25(2):115-137.

Brassard, B.W., and H.Y.H. Chen. 2008. Effects of Forest Type and Disturbance on Diversity of Coarse Woody Debris in Boreal Forest. *Ecosystems* 11(7):1078-1090.

Buongiorno, J., S. Dahir, H.C. Lu, and C.R. Lin. 1994. Tree Size Diversity and Economic Returns in Uneven-Aged Forest Stands. *For. Sci.* 40(1):83-103.

Buongiorno, J., and J.K. Gilles. 2003. Decision methods for forest resource management. Academic Press, San Diego, CA. 439 p.

Cardinale, B.J., D.S. Srivastava, J.E. Duffy, J.P. Wright, A.L. Downing, M. Sankaran, and C. Jouseau. 2006. Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature* 443(7114):989-992.

Caspersen, J.P. 2004. Variation in stand mortality related to successional composition. *For. Ecol. Manage.* 200(1-3):149-160.

Caspersen, J.P., and S.W. Pacala. 2001. Successional diversity and forest ecosystem function. *Ecol. Res.* 16(5):895-903.

Cater, T.C., and F.S. Chapin. 2000. Differential effects of competition or microenvironment on boreal tree seedling establishment after fire. *Ecology* 81(4):1086-1099.

- Chapin, F.S., M.W. Oswood, K. Van Cleve, L.A. Viereck, M.C. Chapin, and D.L. Verbyla (eds.). 2006. *Alaska's changing boreal forest*. Oxford University Press, New York, New York, USA. 368 p.
- Chen, H.Y.H., and R.V. Popadiouk. 2002. Dynamics of North American boreal mixedwoods. *Environmental Reviews* 10(3):137-166.
- Chevan, A., and M. Sutherland. 1991. Hierarchical partitioning. *Am. Stat.* 45(2):90-96.
- Cressie, N.A.C. 1993. *Statistics for spatial data*. John Wiley & Sons, New York, New York, USA.
- Deal, R.L. 2007. Management strategies to increase stand structural diversity and enhance biodiversity in coastal rainforests of Alaska. *Biol. Conserv.* 137(4):520-532.
- Denslow, J.S. 1987. Tropical Rainforest Gaps and Tree Species Diversity. *Annu. Rev. Ecol. Syst.* 18:431-451.
- Don, V.A., W. Arenhövel, R. Jacob, M. Scherer-Lorenzen, and E.D. Schulze. 2007. Anwuchserfolg von 19 verschiedenen Baumarten bei Erstaufforstungen–Ergebnisse eines Biodiversitätsexperiments. *Allg Forst- u J-Ztg* 178:164-172.
- Erskine, P.D., D. Lamb, and M. Bristow. 2006. Tree species diversity and ecosystem function: Can tropical multi-species plantations generate greater productivity? *For. Ecol. Manage.* 233(2-3):205-210.
- Fargione, J., D. Tilman, R. Dybzinski, J. HilleRisLambers, C. Clark, W.S. Harpole, J.M.H. Knops, P.B. Reich, and M. Loreau. 2007. From selection to complementarity: shifts in the causes of biodiversity-productivity relationships in a long-term biodiversity experiment. *Proc. R. Soc. B-Biol. Sci.* 274(1611):871-876.

Folke, C., S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson, and C.S. Holling. 2004. Regime Shifts, Resilience, and Biodiversity in Ecosystem Management. *Annual Review of Ecology, Evolution, and Systematics* 35(ArticleType: research-article / Full publication date: 2004 / Copyright © 2004 Annual Reviews):557-581.

Franklin, J.F. 1981. Ecological characteristics of old growth Douglas-fir forests. USDA Forest Service General Technical Report PNW-118, Pacific Northwest Forest and Range Experiment Station, Portland, Oregon, USA.

Franklin, J.F., T.A. Spies, R. Van Pelt, A.B. Carey, D.A. Thornburgh, D.R. Berg, D.B. Lindenmayer, M.E. Harmon, W.S. Keeton, D.C. Shaw, K. Bible, and J.Q. Chen. 2002. Disturbances and structural development of natural forest ecosystems with silvicultural implications, using Douglas-fir forests as an example. *For. Ecol. Manage.* 155(1-3):399-423.

Gomez-Aparicio, L., R. Zamora, J. Castro, and J.A. Hódar. 2008. Facilitation of tree saplings by nurse plants: Microhabitat amelioration or protection against herbivores? *J. Veg. Sci.* 19(2):161-172.

Gower, S.T., R.E. McMurtrie, and D. Murty. 1996. Aboveground net primary production decline with stand age: Potential causes. *Trends Ecol. Evol.* 11(9):378-382.

Greene, D.F., J.C. Zasada, L. Sirois, D. Kneeshaw, H. Morin, I. Charron, and M.J. Simard. 1999. A review of the regeneration dynamics of North American boreal forest tree species. *Can. J. For. Res* 29(6):824-839.

Grubb, P.J. 1977. The maintenance of species-richness in plant communities: the importance of the regeneration niche. *Biological reviews* 52(1):107-145.

Gutsell, S.L., and E.A. Johnson. 2002. Accurately ageing trees and examining their height-growth rates: implications for interpreting forest dynamics. *J. Ecol.* 90(1):153-166.

Harper, K.A., Y. Bergeron, P. Drapeau, S. Gauthier, and L. De Grandpre. 2006. Changes in spatial pattern of trees and snags during structural development in *Picea mariana* boreal forests. *J. Veg. Sci.* 17(5):625-636.

Johnstone, J., and F. Chapin. 2006. Effects of soil burn severity on post-fire tree recruitment in boreal forest. *Ecosystems* 9(1):14-31.

Johnstone, J.F., F.S. Chapin, III, J. Foote, S. Kemmett, K. Price, and L.A. Viereck. 2004. Decadal observations of tree regeneration following fire in boreal forests. *Canadian Journal of Forest Research* 34(2):267-273.

Johnstone, J.F., T.N. Hollingsworth, F.S. Chapin, and M.C. Mack. 2010. Changes in fire regime break the legacy lock on successional trajectories in Alaskan boreal forest. *Glob. Change Biol.* 16(4):1281-1295.

Jost, L. 2006. Entropy and diversity. *Oikos* 113(2):363-375.

Kneeshaw, D.D., and Y. Bergeron. 1998. Canopy gap characteristics and tree replacement in the southeastern boreal forest. *Ecology* 79(3):783-794.

Legendre, P. 1993. Spatial Autocorrelation - Trouble or New Paradigm. *Ecology* 74(6):1659-1673.

Lei, X.D., W.F. Wang, and C.H. Peng. 2009. Relationships between stand growth and structural diversity in spruce-dominated forests in New Brunswick, Canada. *Can. J. For. Res.-Rev. Can. Rech. For.* 39(10):1835-1847.

Liang, J., J. Buongiorno, and R.A. Monserud. 2005. Growth and yield of all-aged Douglas-fir - western hemlock forest stands: a matrix model with stand diversity effects. *Can. J. For. Res.-Rev. Can. Rech. For.* 35(10):2368-2381.

- Liang, J.J., J. Buongiorno, R.A. Monserud, E.L. Kruger, and M. Zhou. 2007. Effects of diversity of tree species and size on forest basal area growth, recruitment, and mortality. *For. Ecol. Manage.* 243(1):116-127.
- Liang, J.J., and M. Zhou. 2010. A geospatial model of forest dynamics with controlled trend surface. *Ecol. Model.* 221(19):2339-2352.
- Lieffers, V.J., and K.J. Stadt. 1994. Growth of understory *Picea glauca*, *Calamagrostis canadensis*, and *Epilobium angustifolium* in relation to overstory light transmission. *Canadian Journal of Forest Research* 24(6):1193-1198.
- Lindenmayer, D.B., C.R. Margules, and D.B. Botkin. 2000. Indicators of biodiversity for ecologically sustainable forest management. *Conserv. Biol.* 14(4):941-950.
- Loreau, M., and A. Hector. 2001. Partitioning selection and complementarity in biodiversity experiments. *Nature* 412(6842):72-76.
- Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J.P. Grime, A. Hector, D.U. Hooper, M.A. Huston, D. Raffaelli, and B. Schmid. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* 294(5543):804.
- Lusk, C.H., and A. Ortega. 2003. Vertical structure and basal area development in second-growth *Nothofagus* stands in Chile. *J. Appl. Ecol.* 40(4):639-645.
- Mac Nally, R. 1996. Hierarchical partitioning as an interpretative tool in multivariate inference. *Aust. J. Ecol.* 21(2):224-228.
- Mac Nally, R. 2000. Regression and model-building in conservation biology, biogeography and ecology: The distinction between and reconciliation of 'predictive' and 'explanatory' models. *Biodivers. Conserv.* 9(5):655-671.



- Mac Nally, R. 2002. Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables. *Biodivers. Conserv.* 11(8):1397-1401.
- Magurran, A.E. 1988. *Ecological diversity and its measurement*. Princeton University Press, Princeton, N.J., USA.
- Magurran, A.E. 2004. *Measuring biological diversity*. Wiley-Blackwell, Oxford, UK.
- Malone, T., J. Liang, and E.C. Packee. 2009. Cooperative Alaska Forest Inventory. Gen. Tech. Rep. PNW-GTR-785. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 42 p.
- Man, C.D., P.G. Comeau, and D.G. Pitt. 2008. Competitive effects of woody and herbaceous vegetation in a young boreal mixedwood stand. *Can. J. For. Res.-Rev. Can. Rech. For.* 38(7):1817-1828.
- McCarthy, J. 2001. Gap dynamics of forest trees: a review with particular attention to boreal forests. *Environmental Reviews* 9(1):1-59.
- McRoberts, R.E., S. Winter, G. Chirici, E. Hauk, D.R. Pelz, W.K. Moser, and M.A. Hatfield. 2008. Large-scale spatial patterns of forest structural diversity. *Can. J. For. Res.-Rev. Can. Rech. For.* 38(3):429-438.
- Messier, C., R. Doucet, J.C. Ruel, Y. Claveau, C. Kelly, and M.J. Lechowicz. 1999. Functional ecology of advance regeneration in relation to light in boreal forests. *Can. J. For. Res.-Rev. Can. Rech. For.* 29(6):812-823.
- Mladenoff, D. 2005. The promise of landscape modeling: successes, failures, and evolution. P. 90-100 in *Issues and Perspectives in Landscape Ecology*, Wiens, J.A., and M.R. Moss (eds.). Cambridge University Press, Cambridge.

Molofsky, J., and C.K. Augspurger. 1992. The effect of leaf litter on early seedling establishment in a tropical forest. *Ecology* 73(1):68-77.

Mulder, C.P.H., D.D. Uliassi, and D.F. Doak. 2001. Physical stress and diversity-productivity relationships: The role of positive interactions. *Proc. Natl. Acad. Sci. U. S. A.* 98(12):6704-6708.

Naydenov, K., F. Tremblay, Y. Bergeron, and V. Goudiaby. 2006. Influence of active charcoal on the primary growth of jack pine seedlings. *Can. J. For. Res.-Rev. Can. Rech. For.* 36(3):761-767.

Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences* 30(7):683-691.

Pretzsch, H. 2005. Diversity and productivity in forests: evidence from long-term experimental plots. P. 41-64 in *Forest Diversity and Function*, Scherer-Lorenzen, M., C. Körner, and E.-D. Schulze (eds.). Springer-Verlag, Berlin Heidelberg.

Roberts, M.R., and F.S. Gilliam. 1995. Patterns and Mechanisms of Plant Diversity in Forested Ecosystems - Implications for Forest Management. *Ecol. Appl.* 5(4):969-977.

Ruefenacht, B., M.V. Finco, M.D. Nelson, R. Czaplewski, E.H. Helmer, J.A. Blackard, G.R. Holden, A.J. Lister, D. Salajanu, D. Weyermann, and K. Winterberger. 2008. Conterminous US and Alaska Forest Type Mapping Using Forest Inventory and Analysis Data. *Photogramm. Eng. Remote Sens.* 74(11):1379-1388.

Service, U.S.F. 2004. Forest insect and disease conditions in Alaska in 2003. in *U.S. For. Ser. For. Pest Manage. Tech. Rep. R10-TP-123* Anchorage, Alaska.

Shannon, C.E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27:379-423.

Shulski, M., and G. Wendler. 2007. *The climate of Alaska*. University of Alaska Press.

Simpson, E.H. 1949. Measurement of diversity. *Nature* 163(4148):688-688.

Sokal, R., and N. Oden. 1978. Spatial autocorrelation in biology 1. *Methodology. Biological Journal of the Linnean Society* 10(2):199-228.

Stage, A.R., and C. Salas. 2007. Interactions of elevation, aspect, and slope in models of forest species composition and productivity. *For. Sci.* 53(4):486-492.

Team, R.D.C. 2010. *R: A language and environment for statistical computing*. in *R Foundation for Statistical Computing*, Vienna, Austria.

Tilman, D., C.L. Lehman, and K.T. Thomson. 1997. Plant diversity and ecosystem productivity: Theoretical considerations. *Proc. Natl. Acad. Sci. U. S. A.* 94(5):1857-1861.

Tilman, D., P.B. Reich, J. Knops, D. Wedin, T. Mielke, and C. Lehman. 2001. Diversity and productivity in a long-term grassland experiment. *Science* 294(5543):843-845.

Troumbis, A.Y., and D. Memtsas. 2000. Observational evidence that diversity may increase productivity in Mediterranean shrublands. *Oecologia* 125(1):101-108.

Turner, M.G. 2005. Landscape ecology: What is the state of the science? *Annu. Rev. Ecol. Evol. Syst.* 36:319-344.

Ver Hoef, J. 1993. Universal kriging for ecological data. P. 447-453 in *Environmental modeling with GIS*, Goodchild, M.F., B. Parks, and L.T. Steyaert (eds.). Oxford University Press, New York, New York, USA.

Viereck, L., and E. Little. 2007. *Alaska trees and shrubs*. University of Alaska Press, Fairbanks, Alaska, USA.

Vila, M., J. Vayreda, L. Comas, J.J. Ibanez, T. Mata, and B. Obon. 2007. Species richness and wood production: a positive association in Mediterranean forests. *Ecol Lett* 10(3):241-250.

Wagner, H.H., and M.J. Fortin. 2005. Spatial analysis of landscapes: Concepts and statistics. *Ecology* 86(8):1975-1987.

Walsh, C., and R. Mac Nally. 2009. The hier. part package. in R project for statistical computing.

Wardle, D.A., O. Zackrisson, and M.C. Nilsson. 1998. The charcoal effect in Boreal forests: mechanisms and ecological consequences. *Oecologia* 115(3):419-426.

West, G.B., B.J. Enquist, and J.H. Brown. 2009. A general quantitative theory of forest structure and dynamics. *Proc. Natl. Acad. Sci. U. S. A.* 106(17):7040-7045.

Yang, Y.Q., S.M. Huang, S.X. Meng, G. Trincado, and C.L. VanderSchaaf. 2009. A multilevel individual tree basal area increment model for aspen in boreal mixedwood stands. *Can. J. For. Res.-Rev. Can. Rech. For.* 39(11):2203-2214.

Zasada, J.C. 1986. Natural regeneration of trees and tall shrubs on forest sites in interior Alaska. P. 44-73 in *Forest Ecosystems in the Alaskan Taiga: A Synthesis of Structure and Function*, Van Cleve, K., F.S. Chapin III, P.W. Flanagan, L.A. Viereck, and C.T. Dyrness (eds.). Springer-Verlag, New York, New York, USA.

Zasada, J.C., T.L. Sharik, and M. Nygren. 1992. The reproductive process in boreal forest trees. P. 85-125 in *A systems analysis of the global boreal forest*, Shugart, H., R. Leemans, and G. Bonan (eds.). Cambridge University Press, Cambridge, England.

Table 2.1: Frequency of sampled tree species by forest type.

Common name	Scientific name	Frequency (%)			
		Mixed poplar/birch type		Mixed spruce type	
		All live trees	Recruits	All live trees	Recruits
Alaskan Birch	<i>Betula neoalaskana</i>	25.94	2.14	17.24	1.24
Balsam Poplar	<i>Populus balsamifera</i>	2.39	0.40	3.17	0.26
Quaking Aspen	<i>Populus tremuloides</i>	28.62	1.21	9.67	1.77
White Spruce	<i>Picea glauca</i>	31.88	77.35	40.08	55.25
Black Spruce	<i>Picea mariana</i>	8.55	18.63	17.93	35.92
Larch	<i>Larix laricina</i>	0.08	0.13	0.00	0.00
Kenai Birch	<i>Betula kenaica</i>	2.53	0.13	11.91	5.56
All species		100.00	100.00	100.00	100.00

Note: Nomenclature per Viereck and Little (2007).

Table 2.2: Definition, mean, and standard deviation (S.D.) of plot variables used in the analysis

Variable	Description	Unit	Mixed poplar/birch		Mixed spruce	
			Mean	S.D.	Mean	S.D.
Dependent variables						
<i>r</i>	Recruitment, basal area of trees passing the 2.56 cm threshold	m <sup>2</sup> ha <sup>-1</sup> year <sup>-1</sup>	0.02	0.04	0.03	0.05
Stand Diversity Variables						
<i>H<sub>d</sub></i>	Size class basal area diversity		1.88	0.30	1.89	0.31
<i>H<sub>s</sub></i>	Species basal area diversity		0.58	0.31	0.59	0.32
Control Variables						
<i>A</i>	Average stand age	year	71.35	25.79	77.35	31.12
<i>B</i>	Plot basal area in the first inventory	m <sup>2</sup> ha <sup>-1</sup>	23.28	9.42	25.19	10.53
<i>D</i>	Mean diameter of all live trees	cm	14.74	6.10	14.68	7.04
<i>P</i>	Site productivity, derived from stand elevation, slope, and aspect		11.18	30.24	9.62	32.31
<i>L</i>	Percent of plot surface area covered with deadwood of less than 2.5 cm in diameter	percent	48.79	24.40	49.86	22.72
<i>C</i>	Percent of plot surface area covered with charcoal.	percent	9.84	3.10	9.42	3.40
<i>O</i>	Total depth of organic matter	cm	11.42	6.33	10.69	6.19
<i>G</i>	The number of snags (>3 m in height)	ha <sup>-1</sup>	365.9	428.1	457.1	488.1

Table 2.3: Definition, mean, and standard deviation (S.D.) of plot variables used in the analysis continued

$W$	Percent of plot surface area covered with deadwood of greater than 2.5 cm in diameter	percent	13.09	9.15	13.51	8.36
$\lambda$	Easting of UTM coordinates	$10^5$ m	14.47	1.52	13.57	1.38
$\phi$	Northing of UTM coordinates	$10^5$ m	3.36	1.26	3.45	1.38

---

Table 2.4: Percentage contribution (%) to the overall goodness-of-fit and the level of significance (\*P < 0.05; \*\*P < 0.01; \*\*\*P < 0.001) for the explanatory variables of sapling recruitment within mixed poplar/birch and mixed spruce forest types.

	Mixed poplar/birch		Mixed spruce	
$H_d$	12.33	***	12.97	***
$H_s$	5.01	**	7.91	**
A	4.77	*	2.25	
B	4.37		17.95	***
D	29.60	***	41.29	***
P	1.68		0.85	
L	4.46		6.06	
C	7.73	***	0.66	
O	1.60		5.04	
G	26.82	***	3.79	
W	1.63		1.23	



Table 2.5: Difference between the geospatial and non-spatial models of sapling recruitment.

Model / Forest Type	<i>AIC</i>	<i>BIC</i>	<i>L</i>	<i>P</i>
mixed poplar/birch				
a) $0.1014^{***} - 0.0275H_d^{**} + 0.0108H_s$ $+ 0.00003A - 0.0018D^{***} - 0.0013C$ $+ 0.00002G^*$	-811.2	-784.3	413.6	
b) $0.1052^{***} - 0.0219H_d^* + 0.0051H_s$ $+ 0.0001A - 0.0024D^{***} - 0.0019C^*$ $+ 0.00001G^*$	-832.7	-802.5	425.3	<0.001
mixed spruce				
a) $0.0878^{***} - 0.0093H_d + 0.0169H_s$ $- 0.0006B - 0.0023D^{***}$	-731.8	-711.3	371.9	
b) $0.0962^{***} - 0.0140H_d + 0.0059H_s$ $- 0.0005B - 0.0021D^{***}$	-762.9	-735.6	389.5	<0.001

Note:

Level of significance: \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ;

(a) Non-spatial models;

(b) Geospatial models;

*AIC*: Akaike information criterion;

*BIC*: Bayesian information criterion;

*L*: log-likelihood ratio;

*P*: Level of significance for differences between nonspatial and geospatial models.

Table 2.6: Pearson correlations and their levels of significance (\*,  $P < 0.05$ ; \*\*,  $P < 0.01$ ; \*\*\*,  $P < 0.001$ ) between tree size diversity ( $H_d$ ), species diversity ( $H_s$ ), and other explanatory variables for mixed poplar/birch and mixed spruce forest types.

	Mixed poplar/birch		Mixed spruce	
	$H_d$	$H_s$	$H_d$	$H_s$
$H_d$		0.04		-0.03
$H_s$	0.04		-0.03	
$A$	-0.08	-0.11	-0.01	-0.10
$B$	0.53 ***	0.01	0.57 ***	0.08
$D$	0.07	-0.16 *	0.32 ***	-0.12
$P$	-0.12	0.05	0.12	0.06
$L$	0.30 ***	0.01	0.24 **	0.14 *
$C$	-0.06	-0.12	-0.22 **	-0.08
$O$	-0.09	-0.11	-0.24 **	-0.06
$G$	-0.17 *	0.07	-0.03	0.02
$W$	-0.01	-0.13	-0.02	-0.04

Table 2.7: Spatial autocorrelation and its level of significance of linear recruitment model (Eq. 2.3) residuals for mixed poplar/birch and mixed spruce forests types in Alaska.

Forest type	Moran's $I$	$P$ -Value	Geary's $C$	$P$ -Value
Mixed poplar/birch	0.3158	<0.001	0.6843	<0.001
Mixed spruce	0.3805	<0.001	0.5969	<0.001

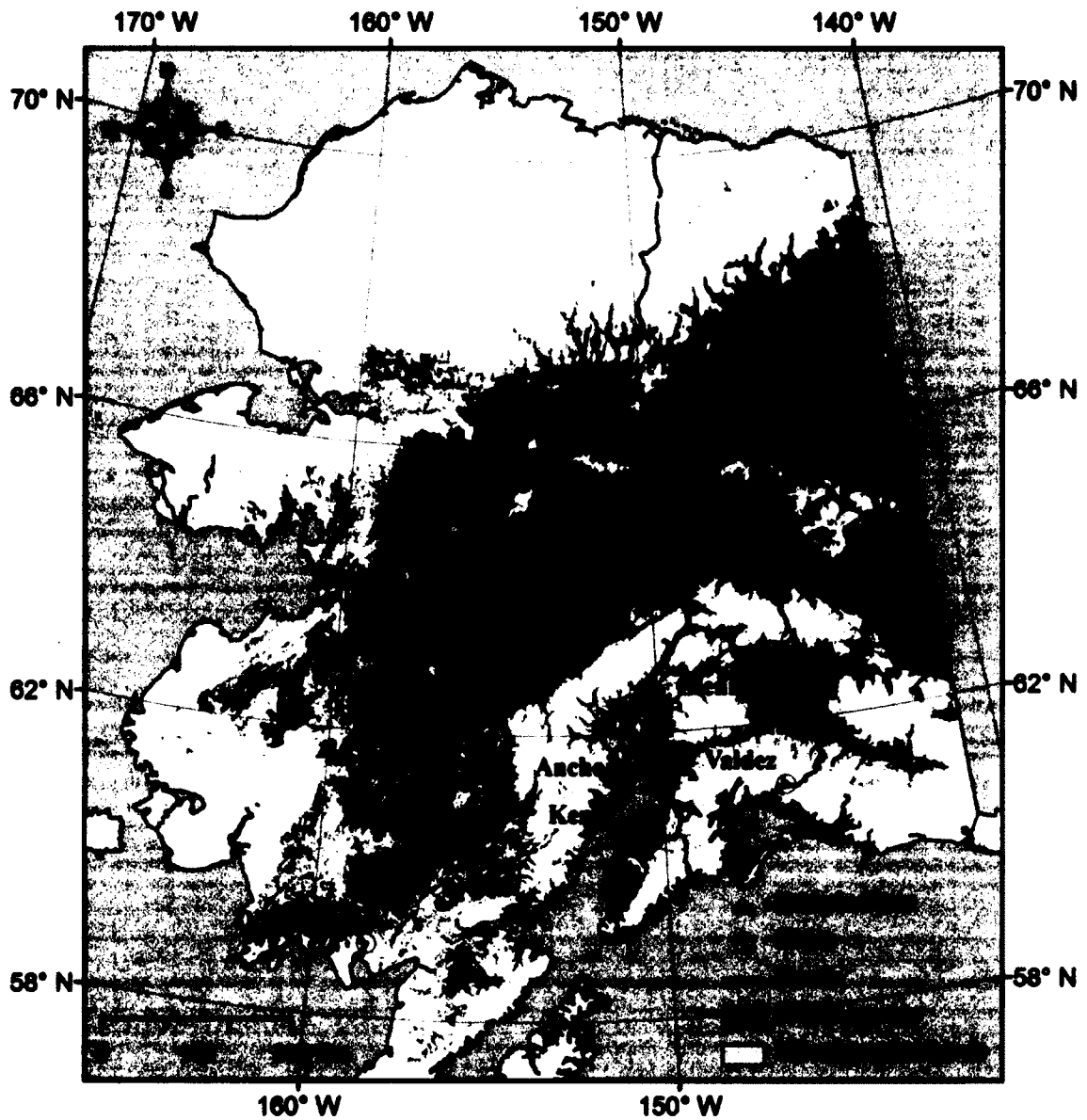


Figure 2.1: Geographic distribution of the 438 Sample Plots (in dots) and the mixed poplar/birch and mixed spruce forest types within the State of Alaska (Ruefenacht et al., 2008).

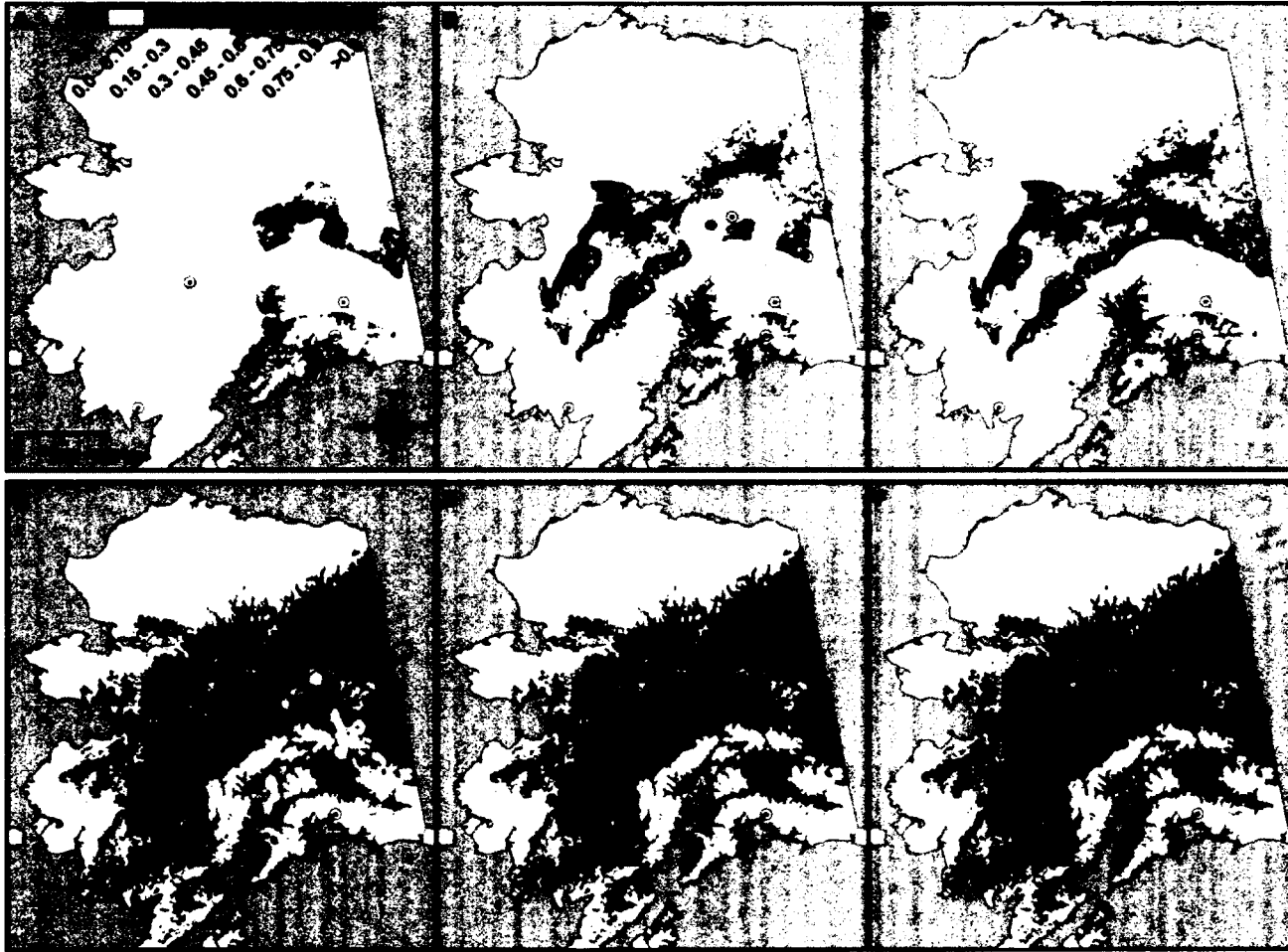


Figure 2.2: Maps of expected sapling recruitment over a 10-year period for the mixed poplar/birch (Figs A, B, and C) and the mixed spruce (Figs D, E, and F) forest types respectively predicted from high ( $H_d = 2.51$ ; Figs A and D), median ( $H_d = 1.89$ ; Figs B and E), and low ( $H_d = 1.27$ ; Figs C and F) tree size diversity within the State of Alaska.

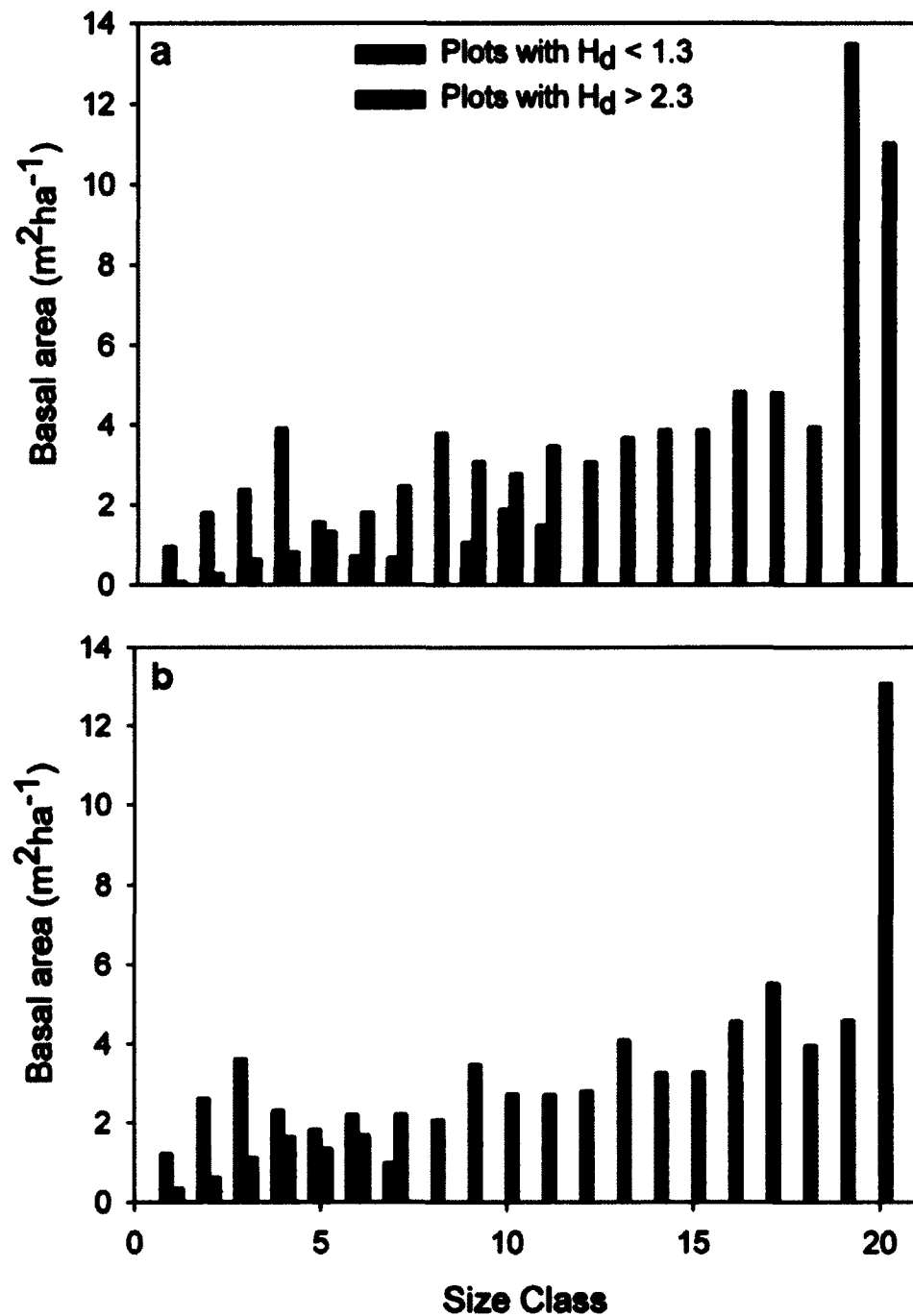


Figure 2.3: Average basal area ( $\text{m}^2\text{ha}^{-1}$ ) of trees on sites with low tree size diversity ( $H_d < 1.3$ ; dark green bars) and high tree size diversity ( $H_d > 2.3$ ; light green bars) within the 20 different size classes (see text for further description) from mixed poplar/birch stands (a) and mixed spruce stands (b).

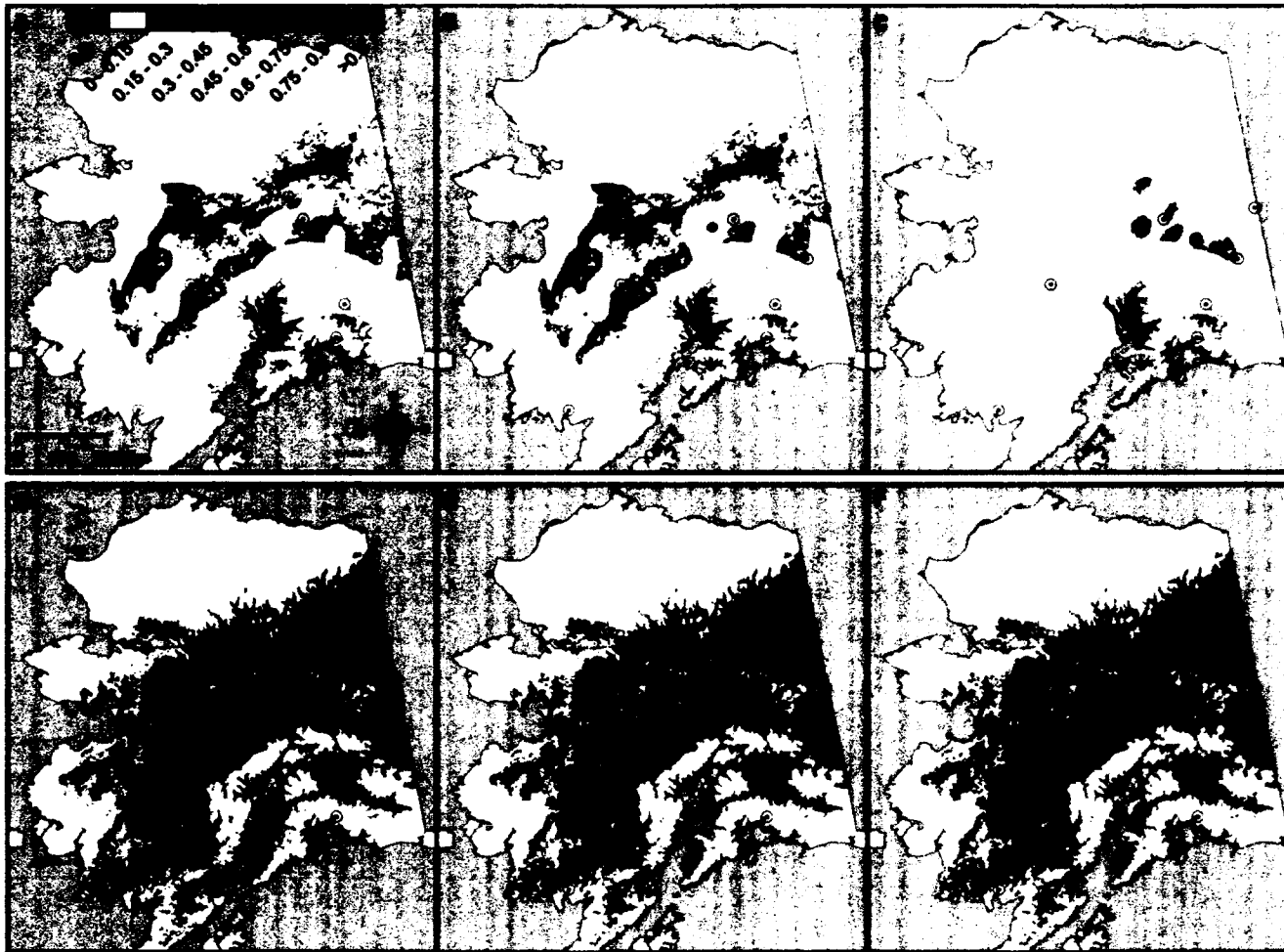


Figure 2.4: Maps of expected sapling recruitment over a 10-year period for the mixed poplar/birch (Figs A, B, and C) and the mixed spruce (Figs D, E, and F) forest types respectively predicted from high ( $H_s = 1.23$ ; Figures A and D), median ( $H_s = 0.59$ ; Figures B and E), and low ( $H_s = 0.00$ ; Figures C and F) species diversity within the State of Alaska.

### Appendix 2.1: R code used in the analysis for Chapter 2 for the popular/birch forest type

**#Load data and Libraries#**

```
popdat <- read.delim(file.choose(),header=T)
```

**#load library**

```
library(geoR)
```

```
library(spdep)
```

```
library(gstat)
```

```
library(sp)
```

```
library(maptools)
```

```
library(RColorBrewer)
```

```
library(rgdal)
```

**#ID Neighborhood structure#**

**#convert coordinates data to matix to be used for neighborhood structure**

```
xy=as.matrix(as.data.frame(cbind(popdat$e,popdat$n)))
```

**#identify neighbors**

```
k1a=knn2nb(knearneigh(xy,k=10))
```

```
all.linked1= max(unlist(nbdists(k1a,xy)))
```

```
dist=
```

```
nb1=dnearneigh(xy, d1=0, d2=dist)
```

```
summary(nb1,xy)
```

**#plot neighbor data**

```
plot(xy)
```

```
plot(nb1,xy, add=TRUE,col="red", lty=2, cex=1)
```

**#convert neighbor matrix to weight matrix**

```
dlist1=nbdists(nb1,xy)
```

```
idlist1=lapply(dlist1,function(x)1/x)
```

```
nbw1=nb2listw(nb1,glist=idlist1,style="W",zero.policy=TRUE)
```

```
summary(nbw1)
```



```

#Data Analysis for Populus.rec#
#convert to geodata for rec
    geopopdat.rec=as.geodata(popdat, coords.col=19:20,data.col=3,covar.col=5:15)
#plot geodata for pop.rec
    plot(geopopdat.rec)
#linear model for pop.rec
    lmpop.rec=lm(Rec~Hdb + Hsb + A + D + C + G, data=popdat)
    summary(lmpop.rec)
#tests for autocorrelation in lm model for pop.rec
    lmrecmoran=moran.mc(lmpop.rec$resid,nbw1,zero.policy=TRUE,nsim=9999)
    lmrecmoran
    lmrecgeary=geary.mc(lmpop.rec$resid,nbw1,zero.policy=TRUE,nsim=9999)
    lmrecgeary
#make variogram for trend model
    va.pop.rec=variog(geopopdat.rec,trend=~Hdb + Hsb + A + D + C + G)
#likfit model for pop.rec
    likpop.rec.trend=likfit(geopopdat.rec,trend=trend.rec,
    ini=c(1,1),cov.model="exponential")
    summary(likpop.rec.trend)
# plot variogram for trend model with likfit model
    plot(va.pop.rec)
    lines(likpop.rec.trend,col="red")
#Setting Kriging Parameters for Birch Aspen #
# define border
    alaskaborder=readShapePoly(file.choose())
#create gstat data format
    coordinates(popdat)=~N+E
#Creating a Birch Aspen distribution grid
    pop.grid=read.dbf(file.choose())

```

```

pop.grid = pop.grid[,3:4]
dim(pop.grid)[1]
n1=dim(pop.grid)[1]
thin=5
select=seq(1,n1,thin)
pop.grid.small=pop.grid[select,]
names(pop.grid.small)=c("x","y")
test.pop = pop.grid.small
plot(test.pop)
points(popdat,pch=19,col="red")
#setting border parms
bbox(alaskaborder)
ll <- c(-600000,950000)
ul <- c(-600000,2350000)
lr <- c(655000,950000)
ur <- c(655000,2350000)
ak.box <- rbind(ll,ul,lr,ur)
ak.box <- as.data.frame(ak.box)
names(ak.box) = c("x","y")
test.pop <- rbind(test.pop,ak.box)
n=dim(test.pop)[1]
# setting explanatory variables at sample means for small prediction grid
Hdb=rep(1.88,n)
Hsb=rep(0.58,n)
A=rep(71.35,n)
B=rep(23.28,n)
D=rep(14.74,n)
P=rep(11.18,n)
L=rep(48.79,n)

```

```

C=rep(9.84,n)
O=rep(11.42,n)
G=rep(365.9,n)
W=rep(13.09,n)
small.grid.mean=cbind(test.pop,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.mean)=~x+y
# explanatory variables at sample means with low Hdb for small prediction grid
Hdb=rep(1.28,n)
Hsb=rep(0.58,n)
small.grid.low=cbind(test.pop,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.low)=~x+y
# explanatory variables at sample means with High Hdb for small prediction grid
Hdb=rep(2.48,n)
Hsb=rep(0.58,n)
small.grid.high=cbind(test.pop,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.high)=~x+y
# explanatory variables at sample means with Low Hsb for small prediction grid
Hdb=rep(1.88,n)
Hsb=rep(0.00,n)
small.grid.low.Hsb=cbind(test.pop,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.low.Hsb)=~x+y
# explanatory variables at sample means with High Hsb for small prediction grid
Hdb=rep(1.88,n)
Hsb=rep(1.20,n)
small.grid.high.Hsb=cbind(test.pop,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.high.Hsb)=~x+y
#Kriging for Populus.rec#
#Variograms for Populus.rec
v.rec= variogram(Rec~Hdb + Hsb + A + D + C + G, popdat)

```

```

v.fit.rec=fit.variogram(v.rec,vgm(0.0010,"Sph",150920,nugget=0.0005))
v.fit.rec
plot(v.rec, model=v.fit.rec,pch=19)
#kriging method using gstat for mean Hdb for truncated grid
krige.rec=krige(Rec~Hdb + Hsb + A + D + C +
G,popdat,small.grid.mean,v.fit.rec)
#kriging for low Hdb
krige.rec.low=krige(Rec~Hdb + Hsb + A + D + C + G,
popdat,small.grid.low,v.fit.rec)
#krige for high Hdb values
krige.rec.high=krige(Rec~Hdb + Hsb + A + D + C +
G,popdat,small.grid.high,v.fit.rec)
#combine all levels of Hdb and plot on one page
rec.high=krige.rec.high[1]
rec.mid=krige.rec[1]
rec.low=krige.rec.low[1]
rec.final=coordinates(rec.high)
year=10
rec.final=cbind(rec.final,rec.high$var1.pred*year)
rec.final=cbind(rec.final,rec.mid$var1.pred*year)
rec.final=cbind(rec.final,rec.low$var1.pred*year)
rec.final=as.data.frame(rec.final)
coordinates(rec.final)=~x+y
names(rec.final)=c("High","Mean","Low")
#export values to ARCGIS
write.csv(rec.final,"poprechd.csv")
#Kriging for Populus.rec.Hsb#
#Variograms for Populus.rec
v.rec= variogram(Rec~Hdb + Hsb + A + D + C + G, popdat)

```

```

v.fit.rec=fit.variogram(v.rec,vgm(0.0010,"Sph",150920,nugget=0.0005))
v.fit.rec
plot(v.rec, model=v.fit.rec,pch=19)
#kriging for mid Hsb
krige.rec.mid.Hsb=krige(Rec~Hdb + Hsb + A + D + C +
G,popdat,small.grid.mean,v.fit.rec)
#krige for low Hsb
krige.rec.low.Hsb=krige(Rec~Hdb + Hsb + A + D + C + G,
popdat,small.grid.low.Hsb,v.fit.rec)
#krige for high Hsb values
krige.rec.high.Hsb=krige(Rec~Hdb + Hsb + A + D + C +
G,popdat,small.grid.high.Hsb,v.fit.rec)
#combine all levels of Hsb and plot on one page
rec.high.Hsb=krige.rec.high.Hsb[1]
rec.mid=krige.rec.mid.Hsb[1]
rec.low.Hsb=krige.rec.low.Hsb[1]
rec.final.Hsb=coordinates(rec.high.Hsb)
year=10
rec.final.Hsb=cbind(rec.final.Hsb,rec.high.Hsb$var1.pred*year)
rec.final.Hsb=cbind(rec.final.Hsb,rec.mid$var1.pred*year)
rec.final.Hsb=cbind(rec.final.Hsb,rec.low.Hsb$var1.pred*year)
rec.final.Hsb=as.data.frame(rec.final.Hsb)
coordinates(rec.final.Hsb)=~x+y
names(rec.final.Hsb)=c("High","Mean","Low")
#export values to ARCGIS
write.csv(rec.final.Hsb,"poprechhs.csv")

```

## Appendix 2.2: R code used in the Analysis for Chapter 2 for the mixed spruce forest type

**#Load data and Libraries#**

```
sprdat <- read.delim(file.choose(),header=T)
```

**#load library**

```
library(geoR)
```

```
library(spdep)
```

```
library(gstat)
```

```
library(sp)
```

```
library(maptools)
```

```
library(RColorBrewer)
```

**#Neighborhood structure#**

**#convert coordinates data to matix to be used for neighborhood structure**

```
xy.spr=as.matrix(as.data.frame(cbind(sprdat$E,sprdat$N)))
```

**#identify neighbors**

```
k.spr=knn2nb(knearneigh(xy.spr,k=10))
```

```
all.linked.spr= max(unlist(nbdists(k.spr,xy.spr)))
```

```
dist.spr=5
```

```
nb.spr=dnearneigh(xy.spr, d1=0, d2=dist.spr)
```

```
summary(nb.spr,xy.spr)
```

**#plot neighbor data**

```
plot(xy.spr)
```

```
plot(nb.spr,xy.spr, add=TRUE,col="red", lty=2, cex=1)
```

**#convert neighbor matrix to weight matrix**

```
dlist.spr=nbdists(nb.spr,xy.spr)
```

```
idlist.spr=lapply(dlist.spr,function(x)1/x)
```

```
nbw.spr=nb2listw(nb.spr,glist=idlist.spr,style="W",zero.policy=TRUE)
```

```
summary(nbw.spr)
```

```

#Data Analysis for Spruce.rec#
#convert to geodata for rec
    geosprdat.rec=as.geodata(sprdat, coords.col=20:21,data.col=3,covar.col=5:15)
#plot geodata for spr.rec
    plot(geosprdat.rec)
#linear model for spr.rec
    lmspr.rec=lm(Rec~Hdb + Hsb + B + D, data=sprdat)
    summary(lmspr.rec)
#tests for autocorrelation in lm model for pop.rec
    lmrecmoran.spr=moran.mc(lmspr.rec$resid,nbw.spr,zero.policy=TRUE,nsim=99
99)
    lmrecmoran.spr
    lmrecgeary.spr=geary.mc(lmspr.rec$resid,nbw.spr,zero.policy=TRUE,nsim=9999
)
    lmrecgeary.spr
#trends for spr.rec
    trend.rec.spr=trend.spatial(~Hdb + Hsb + B + D, sprdat)
# variogram for trend model
    va.spr.rec=variog(geosprdat.mor,trend=~Hdb + Hsb + B + D)
#likfit model for spr.rec
    lkspr.rec.trend=likfit(geosprdat.rec,trend=trend.rec.spr,
    ini=c(1,1),cov.model="spherical")
    summary(lkspr.rec.trend)
#plot variogram for trend model with likfit model
    plot(va.spr.rec)
    lines(lkspr.rec.trend,col="red")
#Setting Kriging Parameters for Spruce#
# define border
    alaskaborder=readShapePoly(file.choose())

```

```

#create gstat data format
  coordinates(sprdat)=~N+E
#Creating a prediction grid
  spr.grid=read.dbf(file.choose())
  spr.grid = spr.grid[,3:4]
  n.1=dim(spr.grid)[1]
  thin.1=15
  select.1=seq(1,n.1,thin.1)
  spr.grid.trunc=spr.grid[select.1,]
  names(spr.grid.trunc)=c("x","y")
  test <- spr.grid.trunc
  plot(test)
  points(sprdat,pch=19,col="red")
#setting border parms
  bbox(alaskaborder)
  ll <- c(-600000,950000)
  ul <- c(-600000,2350000)
  lr <- c(655000,950000)
  ur <- c(655000,2350000)
  ak.box <- rbind(ll,ul,lr,ur)
  ak.box <- as.data.frame(ak.box)
  names(ak.box) = c("x","y")
  test <- rbind(test,ak.box)
  nt <- dim(test)[1]
# setting explanatory variables at sample means
  Hdb=rep(1.89,nt)
  Hsb=rep(0.59,nt)
  A=rep(71.57,nt)
  B=rep(25.19,nt)

```



```

D=rep(14.68,nt)
P=rep(9.62,nt)
L=rep(49.86,nt)
C=rep(9.42,nt)
O=rep(10.69,nt)
G=rep(457.1,nt)
W=rep(13.51,nt)
small.grid.mean.spr=cbind(test,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.mean.spr)=~x+y
# explanatory variables at sample means with low Hdb
Hdb=rep(1.27,nt)
Hsb=rep(0.59,nt)
small.grid.low.spr=cbind(test,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.low.spr)=~x+y
# explanatory variables at sample means with High Hdb
Hdb=rep(2.51,nt)
Hsb=rep(0.59,nt)
small.grid.high.spr=cbind(test,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.high.spr)=~x+y
# explanatory variables at sample means with Low Hsb
Hdb=rep(1.89,nt)
Hsb=rep(0.0,nt)
small.grid.low.Hsb.spr=cbind(test,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.low.Hsb.spr)=~x+y
# explanatory variables at sample means with High Hsb
Hdb=rep(1.88,nt)
Hsb=rep(1.23,nt)
small.grid.high.Hsb.spr=cbind(test,Hdb,Hsb,A,B,D,P,L,C,O,G,W)
coordinates(small.grid.high.Hsb.spr)=~x+y

```

```

#Kriging for Spruce.rec.Hdb#
#Variograms for Spruce.rec
  v.rec.spr= variogram(Rec~Hdb + Hsb + B + D, sprdat)
  v.fit.rec.spr=fit.variogram(v.rec.spr,vgm(0.0011,"Sph",49320,nugget=0.0011))
  v.fit.rec.spr
  plot(v.rec.spr,model=v.fit.rec.spr,pch=19)
#kriging method using gstat for mean Hdb
  krige.rec.spr.t=krige(Rec~Hdb + Hsb + B +
D,sprdat,small.grid.mean.spr,v.fit.rec.spr)
#kriging for low Hdb
  krige.rec.low.spr.t=krige(Rec~Hdb + Hsb + B + D,
sprdat,small.grid.low.spr,v.fit.rec.spr)
#kriging for high Hdb values
  krige.rec.high.spr.t=krige(Rec~Hdb + Hsb + B + D,sprdat, small.grid.high.spr,
v.fit.rec.spr)
#combine all levels of Hdb and plot on one page for truncated grid
  rec.high.spr.t=krige.rec.high.spr.t[1]
  rec.mid.spr.t=krige.rec.spr.t[1]
  rec.low.spr.t=krige.rec.low.spr.t[1]
  rec.final.spr.t=coordinates(rec.high.spr.t)
  rec.final.spr.t=cbind(rec.final.spr.t,rec.high.spr.t$var1.pred*year)
  rec.final.spr.t=cbind(rec.final.spr.t,rec.mid.spr.t$var1.pred*year)
  rec.final.spr.t=cbind(rec.final.spr.t,rec.low.spr.t$var1.pred*year)
  rec.final.spr.t=as.data.frame(rec.final.spr.t)
  coordinates(rec.final.spr.t)=~x+y
  names(rec.final.spr.t)=c("High","Mean","Low")
#setting values to min and max values to adjust scales
  #top left corner point
  rec.final.spr.t$High[1] <- -0.2154

```

```

#top right corner point
rec.final.spr.t$High[dim(rec.final.spr.t)[1]] <- 1.03014
#export values to ARCGIS
write.csv(rec.final.spr.t,"sprrechd.csv")
#Kriging for Spruce.rec.Hsb#
#Variograms for Spruce.rec
v.rec.spr= variogram(Rec~Hdb + Hsb + B + D, sprdat)
v.fit.rec.spr=fit.variogram(v.rec.spr,vgm(0.0011,"Sph",49320,nugget=0.0011))
v.fit.rec.spr
plot(v.rec.spr,model=v.fit.rec.spr,pch=19)
#kriging method using gstat for mean Hsb
krige.rec.spr.t=krige(Rec~Hdb + Hsb + B +
D,sprdat,small.grid.mean.spr,v.fit.rec.spr)
#kriging for low Hsb
krige.rec.low.Hsb.spr.t=krige(Rec~Hdb + Hsb + B + D,
sprdat,small.grid.low.Hsb.spr, v.fit.rec.spr)
#krige for high Hsb values
krige.rec.high.Hsb.spr.t=krige(Rec~Hdb + Hsb + B +
D,sprdat,small.grid.high.Hsb.spr, v.fit.rec.spr)
#combine all levels of Hsb and plot on one page
rec.high.Hsb.spr.t=krige.rec.high.Hsb.spr.t[1]
rec.mid.spr.t=krige.rec.spr.t[1]
rec.low.Hsb.spr.t=krige.rec.low.Hsb.spr.t[1]
rec.final.Hsb.spr.t=coordinates(rec.high.Hsb.spr.t)
rec.final.Hsb.spr.t=cbind(rec.final.Hsb.spr.t,rec.high.Hsb.spr.t$var1.pred*year)
rec.final.Hsb.spr.t=cbind(rec.final.Hsb.spr.t,rec.mid.spr.t$var1.pred*year)
rec.final.Hsb.spr.t=cbind(rec.final.Hsb.spr.t,rec.low.Hsb.spr.t$var1.pred*year)
rec.final.Hsb.spr.t=as.data.frame(rec.final.Hsb.spr.t)
coordinates(rec.final.Hsb.spr.t)=~x+y

```

```
names(rec.final.Hsb.spr.t)=c("High","Mean","Low")
#setting values to min and max values to adjust scales
#top left corner point
rec.final.Hsb.spr.t$High[1] <- -0.1625
#top right corner point
rec.final.Hsb.spr.t$High[dim(rec.final.Hsb.spr.t)[1]] <- 0.9828
#export values to ARCGIS
write.csv(rec.final.Hsb.spr.t,"sprrechs.csv")
```

## Appendix 2.3: Mixed poplar/birch and mixed spruce forest types within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998

### Identification

#### CITATION

##### CITATION INFORMATION

PUBLICATION DATE 2011-04-06

PUBLICATION TIME 000000

#### TITLE

Mixed poplar/birch and mixed spruce forest types within the State of Alaska

GEOSPATIAL DATA PRESENTATION FORM raster digital data

#### SERIES INFORMATION

SERIES NAME Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

ISSUE IDENTIFICATION Forest Ecology and Management 262 (2011) 1608–1617

#### DESCRIPTION

##### ABSTRACT

This raster layer was used in the study that empirically evaluated and mapped the relationship between recruitment and tree size and species diversity (Young et al. 2011) as input to differentiate recruitment between mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. The input forest types were created by Ruefenacht et al. 2008 and constituted forest types across the conterminous US and Alaska. They were downloaded from the website of Oak Ridge National Laboratory Distributed Active Archive Center and clipped to the region of interest.

The region of interest for this study was the Alaskan boreal forest which extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains.

The abstract for the paper in which this data was originally published is below for reference:

Title: Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

Brian Young 1,

Jingjing Liang 1,

**F. Stuart Chapin III 2**

**1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA**

**2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA**

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

**PURPOSE**

**Modeling forest types within the State of Alaska**

**STATUS**

**MAINTENANCE AND UPDATE FREQUENCY   None planned**

**SPATIAL DOMAIN**

**BOUNDING COORDINATES**

**WEST BOUNDING COORDINATE   -162.94478**

**EAST BOUNDING COORDINATE   -137.944993**

**NORTH BOUNDING COORDINATE   68.98776**

**SOUTH BOUNDING COORDINATE   59.094566**

**KEYWORDS**

**THEME**

**THEME KEYWORD THESAURUS   None**

**THEME KEYWORD   Forest types, Landscape heterogeneity**

PLACE  
 PLACE KEYWORD THESAURUS None  
 PLACE KEYWORD Alaska, Boreal forest

ACCESS CONSTRAINTS  
 None

USE CONSTRAINTS  
 None

NATIVE DATA SET ENVIRONMENT  
 Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
 10.0.3.3600

### **Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

RASTER OBJECT INFORMATION  
 RASTER OBJECT TYPE Pixel  
 Row COUNT 21085  
 COLUMN COUNT 22096

### **Spatial Reference**

HORIZONTAL COORDINATE SYSTEM DEFINITION  
 PLANAR  
 MAP PROJECTION  
 MAP PROJECTION NAME NAD 1983 Albers

PLANAR COORDINATE INFORMATION  
 PLANAR COORDINATE ENCODING METHOD coordinate pair  
 COORDINATE REPRESENTATION  
 ABSCISSA RESOLUTION 0.0000000030536018158500163  
 ORDINATE RESOLUTION 0.0000000030536018158500163  
 PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
 HORIZONTAL DATUM NAME D North American 1983  
 ELLIPSOID NAME GRS 1980  
 SEMI-MAJOR AXIS 6378137.0  
 DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Entities and Attributes**

**DETAILED DESCRIPTION****ENTITY TYPE**

**ENTITY TYPE LABEL** forest type.tif.vat

**ATTRIBUTE**

**ATTRIBUTE LABEL** OID

**ATTRIBUTE DEFINITION**

Internal feature number.

**ATTRIBUTE DEFINITION SOURCE** ESRI

**ATTRIBUTE DOMAIN VALUES****UNREPRESENTABLE DOMAIN**

Sequential unique whole numbers that are automatically generated.

**ATTRIBUTE**

**ATTRIBUTE LABEL** VALUE

**ATTRIBUTE DEFINITION**

120 Mixed Spruce

900 Mixed Poplar/ Birch

**ATTRIBUTE**

**ATTRIBUTE LABEL** COUNT

**Metadata Reference**

**METADATA DATE** 2012-08-11

**METADATA CONTACT****CONTACT INFORMATION****CONTACT ORGANIZATION PRIMARY**

**CONTACT ORGANIZATION** University of Alaska Fairbanks

**CONTACT PERSON** Brian Young

**CONTACT ELECTRONIC MAIL ADDRESS** bdyoung@alaska.edu

**METADATA STANDARD NAME** FGDC Content Standard for Digital  
Geospatial Metadata

**METADATA STANDARD VERSION** FGDC-STD-001-1998

**METADATA TIME CONVENTION** local time



**Appendix 2.4: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from high ( $H_d = 2.51$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from high ( $H_d = 2.51$ ) tree size diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

Brian Young 1,

Jingjing Liang 1,

F. Stuart Chapin III 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

Modeling of tree sapling recruitment within the State of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Tree size diversity, Ingrowth, Landscape heterogeneity, Universal kriging

##### **PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska

##### **TEMPORAL**

TEMPORAL KEYWORD THESAURUS None

TEMPORAL KEYWORD Tree species recruitment 1999-2010

ACCESS CONSTRAINTS

None

USE CONSTRAINTS

None

NATIVE DATA SET ENVIRONMENT

Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600

### **Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

RASTER OBJECT INFORMATION

RASTER OBJECT TYPE Pixel

ROW COUNT 255

COLUMN COUNT 251

### **Spatial Reference**

HORIZONTAL COORDINATE SYSTEM DEFINITION

PLANAR

MAP PROJECTION

MAP PROJECTION NAME NAD 1983 Albers

ALBERS CONICAL EQUAL AREA

STANDARD PARALLEL

STANDARD PARALLEL

LONGITUDE OF CENTRAL MERIDIAN

LATITUDE OF PROJECTION ORIGIN

FALSE EASTING

FALSE NORTHING

PLANAR COORDINATE INFORMATION

PLANAR COORDINATE ENCODING METHOD coordinate pair

COORDINATE REPRESENTATION

ABSCISSA RESOLUTION 0.0000000030536018158500163

ORDINATE RESOLUTION 0.0000000030536018158500163

PLANAR DISTANCE UNITS Meter

GEODETIC MODEL

HORIZONTAL DATUM NAME D North American 1983

ELLIPSOID NAME   GRS 1980  
SEMI-MAJOR AXIS   6378137.0  
DENOMINATOR OF FLATTENING RATIO   298.257222101

### **Metadata Reference**

METADATA DATE   2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION   University of Alaska Fairbanks  
CONTACT PERSON   Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS   bdyoung@alaska.edu  
  
METADATA STANDARD NAME   FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION   FGDC-STD-001-1998  
METADATA TIME CONVENTION   local time

**Appendix 2.5: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from median ( $H_d = 1.89$ ) tree size diversity within the State of Alaska**  
**Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from median ( $H_d = 1.89$ ) tree size diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

Brian Young 1,

Jingjing Liang 1,

F. Stuart Chapin III 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK  
99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

Modeling of tree sapling recruitment within the State of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Tree size diversity, Ingrowth, Landscape heterogeneity,  
Universal kriging

##### **PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska

TEMPORAL

TEMPORAL KEYWORD THESAURUS None

TEMPORAL KEYWORD Tree species recruitment 1999-2010

ACCESS CONSTRAINTS

None

USE CONSTRAINTS

None

NATIVE DATA SET ENVIRONMENT

Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600

### **Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

RASTER OBJECT INFORMATION

RASTER OBJECT TYPE Pixel

ROW COUNT 522

COLUMN COUNT 514

### **Spatial Reference**

HORIZONTAL COORDINATE SYSTEM DEFINITION

PLANAR

MAP PROJECTION

MAP PROJECTION NAME NAD 1983 Alaska Albers

ALBERS CONICAL EQUAL AREA

STANDARD PARALLEL 55.0

STANDARD PARALLEL 65.0

LONGITUDE OF CENTRAL MERIDIAN

LATITUDE OF PROJECTION ORIGIN

FALSE EASTING 0.0

FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION

PLANAR COORDINATE ENCODING METHOD coordinate pair

COORDINATE REPRESENTATION

ABSCISSA RESOLUTION 0.0000000030536018158500163

ORDINATE RESOLUTION 0.0000000030536018158500163

PLANAR DISTANCE UNITS   Meter

**GEODETTIC MODEL**

HORIZONTAL DATUM NAME   D North American 1983

ELLIPSOID NAME   GRS 1980

SEMI-MAJOR AXIS   6378137.0

DENOMINATOR OF FLATTENING RATIO   298.257222101

**Metadata Reference**

METADATA DATE   2012-08-11

METADATA CONTACT

CONTACT INFORMATION

CONTACT ORGANIZATION PRIMARY

CONTACT ORGANIZATION   University of Alaska Fairbanks

CONTACT PERSON   Brian Young

CONTACT ELECTRONIC MAIL ADDRESS   bdyoung@alaska.edu

METADATA STANDARD NAME   FGDC Content Standard for Digital  
Geospatial Metadata

METADATA STANDARD VERSION   FGDC-STD-001-1998

METADATA TIME CONVENTION   local time



**Appendix 2.6: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from low ( $H_d = 1.27$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from low ( $H_d = 1.27$ ) tree size diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**Brian Young** 1,

**Jingjing Liang** 1,

**F. Stuart Chapin III** 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK  
99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

**Modeling of tree sapling recruitment within the State of Alaska**

#### **STATUS**

**MAINTENANCE AND UPDATE FREQUENCY** None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

**WEST BOUNDING COORDINATE** -162.94478

**EAST BOUNDING COORDINATE** -137.944993

**NORTH BOUNDING COORDINATE** 68.98776

**SOUTH BOUNDING COORDINATE** 59.094566

#### **KEYWORDS**

##### **THEME**

**THEME KEYWORD THESAURUS** None

**THEME KEYWORD** Tree size diversity, Ingrowth, Landscape heterogeneity, Universal kriging

##### **PLACE**

**PLACE KEYWORD THESAURUS** None

**PLACE KEYWORD** Alaska

TEMPORAL  
 TEMPORAL KEYWORD THESAURUS None  
 TEMPORAL KEYWORD Tree species recruitment 1999-2010

ACCESS CONSTRAINTS  
 None

USE CONSTRAINTS  
 None

NATIVE DATA SET ENVIRONMENT  
 Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
 10.0.3.3600

### **Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

RASTER OBJECT INFORMATION  
 RASTER OBJECT TYPE Pixel  
 ROW COUNT 522  
 COLUMN COUNT 514

### **Spatial Reference**

HORIZONTAL COORDINATE SYSTEM DEFINITION  
 PLANAR  
 MAP PROJECTION  
 MAP PROJECTION NAME NAD 1983 Alaska Albers  
 ALBERS CONICAL EQUAL AREA  
 STANDARD PARALLEL 55.0  
 STANDARD PARALLEL 65.0  
 LONGITUDE OF CENTRAL MERIDIAN  
 LATITUDE OF PROJECTION ORIGIN  
 FALSE EASTING 0.0  
 FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION  
 PLANAR COORDINATE ENCODING METHOD coordinate pair  
 COORDINATE REPRESENTATION  
 ABSCISSA RESOLUTION 0.0000000030536018158500163  
 ORDINATE RESOLUTION 0.0000000030536018158500163  
 PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time

**Appendix 2.7: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from high ( $H_d = 2.51$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from high ( $H_d = 2.51$ ) tree size diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree size diversity, as measured with the Shannon's index, within mixed spruce stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

Brian Young 1,

Jingjing Liang 1,

F. Stuart Chapin III 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

Modeling of tree sapling recruitment within the State of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Tree size diversity, Ingrowth, Landscape heterogeneity, Universal kriging

##### **PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska

**TEMPORAL****TEMPORAL KEYWORD THESAURUS** None**TEMPORAL KEYWORD** Tree species recruitment 1999-2010**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600**Spatial Data Organization****DIRECT SPATIAL REFERENCE METHOD** Raster**RASTER OBJECT INFORMATION****RASTER OBJECT TYPE** Pixel**ROW COUNT** 332**COLUMN COUNT** 295**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION****PLANAR****MAP PROJECTION****MAP PROJECTION NAME** NAD 1983 Albers**ALBERS CONICAL EQUAL AREA****STANDARD PARALLEL****STANDARD PARALLEL****LONGITUDE OF CENTRAL MERIDIAN****LATITUDE OF PROJECTION ORIGIN****FALSE EASTING****FALSE NORTHING****PLANAR COORDINATE INFORMATION****PLANAR COORDINATE ENCODING METHOD** coordinate pair**COORDINATE REPRESENTATION****ABSCISSA RESOLUTION** 0.0000000030536018158500163**ORDINATE RESOLUTION** 0.0000000030536018158500163**PLANAR DISTANCE UNITS** Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time



**Appendix 2.8: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from median ( $H_d = 1.89$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from median ( $H_d = 1.89$ ) tree size diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree size diversity, as measured with the Shannon's index, within mixed spruce stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

Brian Young 1,

Jingjing Liang 1,

F. Stuart Chapin III 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK  
99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

**Modeling of tree sapling recruitment within the State of Alaska**

#### **STATUS**

**MAINTENANCE AND UPDATE FREQUENCY** None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

**WEST BOUNDING COORDINATE** -162.94478

**EAST BOUNDING COORDINATE** -137.944993

**NORTH BOUNDING COORDINATE** 68.98776

**SOUTH BOUNDING COORDINATE** 59.094566

#### **KEYWORDS**

##### **THEME**

**THEME KEYWORD THESAURUS** None

**THEME KEYWORD** Tree size diversity, Ingrowth, Landscape heterogeneity, Universal kriging

##### **PLACE**

**PLACE KEYWORD THESAURUS** None

**PLACE KEYWORD** Alaska

TEMPORAL  
 TEMPORAL KEYWORD THESAURUS None  
 TEMPORAL KEYWORD Tree species recruitment 1999-2010

ACCESS CONSTRAINTS  
 None

USE CONSTRAINTS  
 None

NATIVE DATA SET ENVIRONMENT  
 Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
 10.0.3.3600

### **Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

RASTER OBJECT INFORMATION  
 RASTER OBJECT TYPE Pixel  
 ROW COUNT 679  
 COLUMN COUNT 603

### **Spatial Reference**

HORIZONTAL COORDINATE SYSTEM DEFINITION  
 PLANAR  
 MAP PROJECTION  
 MAP PROJECTION NAME NAD 1983 Albers  
 ALBERS CONICAL EQUAL AREA  
 STANDARD PARALLEL  
 STANDARD PARALLEL  
 LONGITUDE OF CENTRAL MERIDIAN  
 LATITUDE OF PROJECTION ORIGIN  
 FALSE EASTING  
 FALSE NORTHING

PLANAR COORDINATE INFORMATION  
 PLANAR COORDINATE ENCODING METHOD coordinate pair  
 COORDINATE REPRESENTATION  
 ABSCISSA RESOLUTION 0.0000000030536018158500163  
 ORDINATE RESOLUTION 0.0000000030536018158500163  
 PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time

**Appendix 2.9: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from low ( $H_d = 1.27$ ) tree size diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from low ( $H_d = 1.27$ ) tree size diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree size diversity, as measured with the Shannon's index, within mixed spruce stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

Brian Young 1,

Jingjing Liang 1,

F. Stuart Chapin III 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK  
99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

**Modeling of tree sapling recruitment within the State of Alaska**

#### **STATUS**

**MAINTENANCE AND UPDATE FREQUENCY** None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

**WEST BOUNDING COORDINATE** -162.94478

**EAST BOUNDING COORDINATE** -137.944993

**NORTH BOUNDING COORDINATE** 68.98776

**SOUTH BOUNDING COORDINATE** 59.094566

#### **KEYWORDS**

##### **THEME**

**THEME KEYWORD THESAURUS** None

**THEME KEYWORD** Tree size diversity, Ingrowth, Landscape heterogeneity, Universal kriging

##### **PLACE**

**PLACE KEYWORD THESAURUS** None

**PLACE KEYWORD** Alaska

**TEMPORAL****TEMPORAL KEYWORD THESAURUS** None**TEMPORAL KEYWORD** Tree species recruitment 1999-2010**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600**Spatial Data Organization****DIRECT SPATIAL REFERENCE METHOD** Raster**RASTER OBJECT INFORMATION****RASTER OBJECT TYPE** Pixel**ROW COUNT** 679**COLUMN COUNT** 603**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION****PLANAR****MAP PROJECTION****MAP PROJECTION NAME** NAD 1983 Albers**ALBERS CONICAL EQUAL AREA****STANDARD PARALLEL****STANDARD PARALLEL****LONGITUDE OF CENTRAL MERIDIAN****LATITUDE OF PROJECTION ORIGIN****FALSE EASTING****FALSE NORTHING****PLANAR COORDINATE INFORMATION****PLANAR COORDINATE ENCODING METHOD** coordinate pair**COORDINATE REPRESENTATION****ABSCISSA RESOLUTION** 0.0000000030536018158500163**ORDINATE RESOLUTION** 0.0000000030536018158500163**PLANAR DISTANCE UNITS** Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time



**Appendix 2.10: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from high ( $H_s = 1.23$ ) tree species diversity within the State of Alaska**  
**Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from high ( $H_s = 1.23$ ) tree species diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree species diversity, as measured with the Shannon's index, within mixed poplar/birch stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**Brian Young** 1,

**Jingjing Liang** 1,

**F. Stuart Chapin III** 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK  
99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

Modeling of tree sapling recruitment within the State of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.944780

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.987760

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Species diversity, Ingrowth, Landscape heterogeneity,  
Universal kriging

##### **PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska

**TEMPORAL**

TEMPORAL KEYWORD THESAURUS None

TEMPORAL KEYWORD Tree species recruitment 1999-2010

**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600**Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

**RASTER OBJECT INFORMATION**

RASTER OBJECT TYPE Pixel

ROW COUNT 522

COLUMN COUNT 514

**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION**

PLANAR

MAP PROJECTION

MAP PROJECTION NAME NAD 1983 Alaska Albers

ALBERS CONICAL EQUAL AREA

STANDARD PARALLEL 55.0

STANDARD PARALLEL 65.0

LONGITUDE OF CENTRAL MERIDIAN

LATITUDE OF PROJECTION ORIGIN

FALSE EASTING 0.0

FALSE NORTHING 0.0

**PLANAR COORDINATE INFORMATION**

PLANAR COORDINATE ENCODING METHOD coordinate pair

COORDINATE REPRESENTATION

ABSCISSA RESOLUTION 0.0000000030536018158500163

ORDINATE RESOLUTION 0.0000000030536018158500163

PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time

**Appendix 2.11: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from median ( $H_s = 0.59$ ) tree species diversity within the State of Alaska**

**Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from median ( $H_s = 0.59$ ) tree species diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree species diversity, as measured with the Shannon's index, within mixed poplar/birch stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**Brian Young** 1,

**Jingjing Liang** 1,

**F. Stuart Chapin III** 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK  
99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

Modeling of tree sapling recruitment within the State of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Species diversity, Ingrowth, Landscape heterogeneity,  
Universal kriging

##### **PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska

**TEMPORAL****TEMPORAL KEYWORD THESAURUS** None**TEMPORAL KEYWORD** Tree species recruitment 1999-2010**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600**Spatial Data Organization****DIRECT SPATIAL REFERENCE METHOD** Raster**RASTER OBJECT INFORMATION****RASTER OBJECT TYPE** Pixel**ROW COUNT** 255**COLUMN COUNT** 251**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION****PLANAR****MAP PROJECTION****MAP PROJECTION NAME** NAD 1983 Albers**ALBERS CONICAL EQUAL AREA****STANDARD PARALLEL****STANDARD PARALLEL****LONGITUDE OF CENTRAL MERIDIAN****LATITUDE OF PROJECTION ORIGIN****FALSE EASTING****FALSE NORTHING****PLANAR COORDINATE INFORMATION****PLANAR COORDINATE ENCODING METHOD** coordinate pair**COORDINATE REPRESENTATION****ABSCISSA RESOLUTION** 0.0000000030536018158500163**ORDINATE RESOLUTION** 0.0000000030536018158500163**PLANAR DISTANCE UNITS** Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time



**Appendix 2.12: Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from low ( $H_s = 0.00$ ) tree species diversity within the State of Alaska**

**Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed poplar/birch forest type predicted from low ( $H_s = 0.00$ ) tree species diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree species diversity, as measured with the Shannon's index, within mixed poplar/birch stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

Title: Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

Brian Young 1,

Jingjing Liang 1,

F. Stuart Chapin III 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree species diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

Modeling of tree sapling recruitment within the State of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Species diversity, Ingrowth, Landscape heterogeneity, Universal kriging

##### **PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska

TEMPORAL  
 TEMPORAL KEYWORD THESAURUS None  
 TEMPORAL KEYWORD Tree species recruitment 1999-2010

ACCESS CONSTRAINTS  
 None

USE CONSTRAINTS  
 None

NATIVE DATA SET ENVIRONMENT  
 Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
 10.0.3.3600

### **Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

RASTER OBJECT INFORMATION  
 RASTER OBJECT TYPE Pixel  
 ROW COUNT 522  
 COLUMN COUNT 514

### **Spatial Reference**

HORIZONTAL COORDINATE SYSTEM DEFINITION  
 PLANAR  
 MAP PROJECTION  
 MAP PROJECTION NAME NAD 1983 Alaska Albers  
 ALBERS CONICAL EQUAL AREA  
 STANDARD PARALLEL 55.0  
 STANDARD PARALLEL 65.0  
 LONGITUDE OF CENTRAL MERIDIAN  
 LATITUDE OF PROJECTION ORIGIN  
 FALSE EASTING 0.0  
 FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION  
 PLANAR COORDINATE ENCODING METHOD coordinate pair  
 COORDINATE REPRESENTATION  
 ABSCISSA RESOLUTION 0.0000000030536018158500163  
 ORDINATE RESOLUTION 0.0000000030536018158500163  
 PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time

**Appendix 2.13: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from high ( $H_s = 1.23$ ) tree species diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from high ( $H_s = 1.23$ ) tree species diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree species diversity, as measured with the Shannon's index, within mixed spruce stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**Brian Young** 1,

**Jingjing Liang** 1,

**F. Stuart Chapin III** 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK  
99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

Modeling of tree sapling recruitment within the State of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Species diversity, Ingrowth, Landscape heterogeneity,  
Universal kriging

##### **PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska

**TEMPORAL**

TEMPORAL KEYWORD THESAURUS None

TEMPORAL KEYWORD Tree species recruitment 1999-2010

**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600**Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

**RASTER OBJECT INFORMATION**

RASTER OBJECT TYPE Pixel

ROW COUNT 283

COLUMN COUNT 252

**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION**

PLANAR

MAP PROJECTION

MAP PROJECTION NAME NAD 1983 Albers

ALBERS CONICAL EQUAL AREA

STANDARD PARALLEL

STANDARD PARALLEL

LONGITUDE OF CENTRAL MERIDIAN

LATITUDE OF PROJECTION ORIGIN

FALSE EASTING

FALSE NORTHING

**PLANAR COORDINATE INFORMATION**

PLANAR COORDINATE ENCODING METHOD coordinate pair

COORDINATE REPRESENTATION

ABSCISSA RESOLUTION 0.0000000030536018158500163

ORDINATE RESOLUTION 0.0000000030536018158500163

PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time



**Appendix 2.14: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from median ( $H_s = 0.59$ ) tree species diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from median ( $H_s = 0.59$ ) tree species diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree species diversity, as measured with the Shannon's index, within mixed spruce stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

Brian Young 1,

Jingjing Liang 1,

F. Stuart Chapin III 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA  
 2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

Modeling of tree sapling recruitment within the State of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Species diversity, Ingrowth, Landscape heterogeneity, Universal kriging

##### **PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska

TEMPORAL

TEMPORAL KEYWORD THESAURUS None

TEMPORAL KEYWORD Tree species recruitment 1999-2010

ACCESS CONSTRAINTS

None

USE CONSTRAINTS

None

NATIVE DATA SET ENVIRONMENT

Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600

### **Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

RASTER OBJECT INFORMATION

RASTER OBJECT TYPE Pixel

ROW COUNT 679

COLUMN COUNT 603

### **Spatial Reference**

HORIZONTAL COORDINATE SYSTEM DEFINITION

PLANAR

MAP PROJECTION

MAP PROJECTION NAME NAD 1983 Albers

ALBERS CONICAL EQUAL AREA

STANDARD PARALLEL

STANDARD PARALLEL

LONGITUDE OF CENTRAL MERIDIAN

LATITUDE OF PROJECTION ORIGIN

FALSE EASTING

FALSE NORTHING

PLANAR COORDINATE INFORMATION

PLANAR COORDINATE ENCODING METHOD coordinate pair

COORDINATE REPRESENTATION

ABSCISSA RESOLUTION 0.0000000030536018158500163

ORDINATE RESOLUTION 0.0000000030536018158500163

PLANAR DISTANCE UNITS Meter

**GEODETTIC MODEL**

HORIZONTAL DATUM NAME D North American 1983

ELLIPSOID NAME GRS 1980

SEMI-MAJOR AXIS 6378137.0

DENOMINATOR OF FLATTENING RATIO 298.257222101

**Metadata Reference**

METADATA DATE 2012-08-11

METADATA CONTACT

CONTACT INFORMATION

CONTACT ORGANIZATION PRIMARY

CONTACT ORGANIZATION University of Alaska Fairbanks

CONTACT PERSON Brian Young

CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu

METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata

METADATA STANDARD VERSION FGDC-STD-001-1998

METADATA TIME CONVENTION local time

**Appendix 2.15: Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from low ( $H_s = 0.00$ ) tree species diversity within the State of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Sapling recruitment over a 10-year period for a mixed spruce forest type predicted from low ( $H_s = 0.00$ ) tree species diversity within the State of Alaska**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**ISSUE IDENTIFICATION** Forest Ecology and Management 262 (2011) 1608–1617

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that empirically evaluated and mapped the relationship between recruitment and tree species diversity, as measured with the Shannon's index, within mixed spruce stands across the boreal forest of Alaska. The effects of tree diversity on recruitment were then projected over a 10 year time period.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach

**Brian Young** 1,

**Jingjing Liang** 1,

**F. Stuart Chapin III** 2

1 Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2 Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK  
99775-7000, USA

**Abstract:** This study empirically evaluates and maps the relationships between recruitment and species and tree size diversity, as measured with the Shannon's index, within mixed poplar/birch and mixed spruce stands across the boreal forest of Alaska. Data were collected from 438 permanent sample plots re-measured at a 5-year interval. Significant explanatory factors of recruitment, including species and tree size diversity were first identified using hierarchical partitioning. The effects of tree diversity on recruitment were then studied using generalized linear models and universal kriging to account for non-spatial factors and for spatial autocorrelation. We found a consistent positive relationship between recruitment and species diversity and a general negative relationship between recruitment and tree size diversity, indicating a tradeoff between species diversity and tree size diversity in affecting recruitment. These relationships however were not uniform across the landscape, presumably because they were subject to strong spatial autocorrelation attributable to natural disturbances and environmental stressors. In general, diversity had least effect on recruitment in stressful environments where stress, rather than competition, most likely governed recruitment.

#### **PURPOSE**

Modeling of tree sapling recruitment within the State of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Species diversity, Ingrowth, Landscape heterogeneity,  
Universal kriging

##### **PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska

**TEMPORAL****TEMPORAL KEYWORD THESAURUS** None**TEMPORAL KEYWORD** Tree species recruitment 1999-2010**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600**Spatial Data Organization****DIRECT SPATIAL REFERENCE METHOD** Raster**RASTER OBJECT INFORMATION****RASTER OBJECT TYPE** Pixel**ROW COUNT** 679**COLUMN COUNT** 603**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION****PLANAR****MAP PROJECTION****MAP PROJECTION NAME** NAD 1983 Albers**ALBERS CONICAL EQUAL AREA****STANDARD PARALLEL****STANDARD PARALLEL****LONGITUDE OF CENTRAL MERIDIAN****LATITUDE OF PROJECTION ORIGIN****FALSE EASTING****FALSE NORTHING****PLANAR COORDINATE INFORMATION****PLANAR COORDINATE ENCODING METHOD** coordinate pair**COORDINATE REPRESENTATION****ABSCISSA RESOLUTION** 0.0000000030536018158500163**ORDINATE RESOLUTION** 0.0000000030536018158500163**PLANAR DISTANCE UNITS** Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-11  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time



## CHAPTER 3: MODELING AND MAPPING FOREST DIVERSITY OF INTERIOR ALASKA AT 1-KM<sup>2</sup> RESOLUTION FOR CURRENT AND POSSIBLE FUTURE CLIMATE CONDITIONS<sup>3.1</sup>

### 3.1 ABSTRACT

Proactive forest planning requires spatially accurate information about forest diversity. The most cost-efficient way to obtain this information is through modeling, i.e., predicting key forest diversity measures as a function of environmental factors. Patterns of forest diversity are less well known in the boreal forest of interior Alaska compared to most ecosystems of North America. To understand the diversity patterns of this forest, we employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables. We developed prediction maps for the current levels of tree size-class and species diversity and created maps showing the potential changes to these values under a null climate change scenario and for the International Panel on Climate Change (IPCC) “mid-range” scenario (A1B) for the year 2050. The method employed here yielded good accuracy for the huge Alaskan landscape despite the exclusion of spectral reflectance data due to its transient nature. The results indicate that the geographic pattern of tree species diversity differs from the pattern of tree size-class diversity across this forest type and that future climate scenarios have different effects on tree species and tree size-class diversity depending on location. The results also suggest that human impact factors have a greater effect than the ecological factors in predicting the patterns of diversity within the boreal forest of interior Alaska.

---

<sup>3.1</sup> Submitted in a slightly modified form to *Landscape Ecology* as: Young B, Yarie J, Verbyla D, Huettmann F, Chapin FS. Modeling and mapping forest diversity of interior Alaska at 1-km<sup>2</sup> resolution for current and possible future climate conditions.

### 3.2 INTRODUCTION

Maintaining or even potentially enhancing biodiversity across the landscape has both political and ecological importance (Rands et al. 2010). Plant biodiversity is commonly measured using species richness, which is an important (Condit et al. 2006; Hubbell 2001), although not the only measure of biodiversity. The variation among species in plant sizes, i.e., the structural diversity, also contributes to the overall diversity of an area (Franklin 1988; Tilman et al. 1997). Biodiversity is scale-dependent (Crawley and Harral 2001; Magurran 1988; Wiens 1989). Several studies have analyzed biodiversity at both large and small scales (see for instance Hooper et al. 2005). However, less is known at the intermediate  $\sim 0.1 \text{ km}^2$  to  $3 \text{ km}^2$ , or mesoscale (Crawley and Harral 2001; Heikkinen 1996; Niemelä 1999), which is typically the scale at which forestry decisions are made (Niemelä 1999; O'Neill et al. 1997).

Forest inventory data combined with remote sensing data from both satellite and aerial platforms allow for the mapping of various forest parameters at the landscape scale (see for instance Fassnacht et al. 2006; Iverson and Prasad 2001; McRoberts et al. 2008; Ruefenacht et al. 2008). Most studies used inventory data that were uniformly distributed across the landscape. In regions with sparse or non-uniformly distributed inventory data, forest-inventory and remote-sensing data have been combined using spatial interpolation techniques to predict the geographical distribution of forest attributes (Liang and Zhou 2010; Parmentier et al. 2011; Young et al. 2011). Combining forest inventory and remotely sensed data in isolated regions to determine landscape level patterns of biodiversity at the mesoscale has, however, received less attention.

Maps depicting the spatially explicit patterns of biodiversity are valuable for planning and monitoring (Drew et al. 2010). Creating such maps typically employs predictive spatial modeling techniques where the parameters of interest are obtained from inventory data and then related to remotely mapped attributes (see Austin 2002; Cushman and Huettmann 2010; Cushman and McKelvey 2009; Ferrier et al. 2002; Franklin 1995; Guisan and Thuiller 2005). Several statistical approaches have been used to create such maps, with non-parametric approaches tending to yield better results than parametric

approaches, which often violate statistical assumptions (Drew et al. 2010; Prasad et al. 2006). Spatial autocorrelation, a general property of most ecological attributes (Bivand et al. 2008; Legendre 1993), is an additional issue to address, especially in large-scale forest studies (Liang and Zhou 2010). When unaccounted for, spatial autocorrelation may affect statistical model predictions because it violates the assumption of independence on which most standard statistical procedures rely (Legendre 1993). Thus, non-parametric models that account for, or are tolerant of, spatial autocorrelation and noisy data could be generally useful in assessing the spatial patterns of biodiversity (Craig and Huettmann 2008; Li et al. 2011).

Several indices of biodiversity have been applied to forested ecosystems (e.g., Lei et al. 2009; Liang et al. 2007; Lindenmayer et al. 2000; McRoberts et al. 2008). One of the most common indices is the Shannon's index (Shannon 1948), which reflects both evenness and richness of species by weighing all species proportionately to their frequencies in the sample (Jost 2006; Magurran 1988). This index has been used effectively to measure forest stand diversity (Buongiorno et al. 1994; Liang et al. 2006, 2007; Varga et al. 2005; Young et al. 2011) and outperforms other indices in boreal forests (Young et al. 2011). We therefore, chose Shannon's index as the basis for our diversity measure in this study.

Estimation of biodiversity across broad areas requires predictive variables that can be easily analyzed and correlated with diversity measures. These variables should maintain their bias, accuracy, and precision as far as possible when applied at different scales (Hellmann and Fowler 1999). Patterns of biodiversity are strongly influenced by the scale at which it is assessed (Crawley and Harral 2001; Wiens 1989). A top-down approach has been viewed as the best approach when considering global factors, such as climate, are predictive variables (Whittaker et al. 2001). This approach, however, may yield potential conflicts due to trade-offs between the scale of species occurrences and the scale of management (Bunnell and Huggard 1999). An intermediate mesoscale level of analysis has the ability to detect both large-scale landscape patterns and local patch

dynamics (Bunnell and Huggard 1999; Whittaker et al. 2001) and was therefore the scale used in this study.

Forests change dynamically due to successional processes. However, climate change complicates our ability to predict how trees and forest ecosystems will respond in the future. A number of ecological models have combined forest dynamics and climate change scenarios to predict future stand states (e.g., Iverson and Prasad 2001; Nightingale et al. 2008; Prasad et al. 2006). However, the scale at which these models were applied was rather large ( $\geq 20$ -km cell). To that end, the results of these prior modeling attempts may not be as useful for land-use planning because forestry operations typically occur at the mesoscale (Kneeshaw et al. 2000; Niemelä 1999).

The boreal forest of Alaska extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains (Fig.3.1), covering an area of nearly 500,000 km<sup>2</sup>. The Alaskan boreal forest consists of a mosaic of two general forest types, mixed poplar/birch and mixed spruce (Ruefenacht et al. 2008; Viereck and Little 2007; Young et al. 2011) and primarily contains seven tree species. White spruce (*Picea glauca*) and black spruce (*P. marianana*) are the predominant conifers and two poplars (*Populus tremuloides* and *P. balsamifera*), two birches (*Betula neoalaska* and *B. kенаica*), and tamarack larch (*Larix laricina*) represent the deciduous species. Despite the relative floristic simplicity, the boreal forest is the result of complex interactions between climate, topography, geology, species ecology, and disturbances (see Chapin et al. 2006c).

Within interior Alaska, significant variation in tree growth occurs due to local differences in topography, soil type, the biota, the successional state, and climate conditions (Chapin et al. 2006b; Liang 2010; Lloyd and Fastie 2002; Van Cleve et al. 1983; Wilmsking and Juday 2005). These variables are collectively referred to as state factors (Major 1951) and have been linked to ecosystem function within the boreal forest (Chapin et al. 2006b; Hollingsworth et al. 2010; Yarie and Van Cleve 2010). A change in state factor can significantly influence forest diversity (Chapin et al. 2006b). For

example, a longer, warmer growing season, as predicted in some future climate models (Walsh et al. 2008), may preferentially benefit the growth of some tree species over others (Juday et al. 2005; McGuire et al. 2010). Additionally, fire regimes within the boreal forest appear to be changing and are projected to continue to do so in the future (Kasischke et al. 2010). This expected change may alter forest successional processes, resulting in changes in the biotic makeup of the forest (Johnstone et al. 2010; Serreze et al. 2000).

Here we analyze forest inventory data using machine learning to investigate the current tree species and tree size-class diversity within the boreal forest of Alaska and to predict possible future scenarios for these two diversity measures. Using this method, we developed a spatially dynamic model depicting forest diversity for the Alaskan boreal forest. With remote sensing data and the geographic information system (GIS) tools, stand-level predictions were aggregated to map forest diversity for the entire region.

### 3.3 METHODS AND MATERIALS

#### 3.3.1 DIVERSITY DATA

Our dataset consisted of 704 permanent sample plots (PSPs) from the Cooperative Alaska Forest Inventory Database (CAFI; <http://www.lter.uaf.edu/data>) (Malone et al. 2009) and the Fort Wainwright Forest Inventory Database (WAIN) (Rees, personal communication). These databases consist of periodically re-measured permanent sample plots (PSPs) located across interior and south-central Alaska north of 60°N (Fig. 1). These PSPs are primarily located on well-stocked forested lands. The CAFI plots are primarily located along the road system on Federal, State, Borough, and Native Corporation lands, while the WAIN plots are scattered across Military lands (Figure 3.1).

The forest diversity on each of the 704 PSPs was calculated using Shannon's index (Shannon 1948) to measure tree species diversity,  $H_s$ , and tree size-class diversity,  $H_d$  on each plot (Table 3.1):

$$H_s = -\sum_{i=1}^{ns} \frac{B_i}{B} \cdot \ln\left(\frac{B_i}{B}\right) \quad (3.1)$$

and

$$H_d = -\sum_{j=1}^{nd} \frac{B_j}{B} \cdot \ln\left(\frac{B_j}{B}\right) \quad (3.2)$$

where  $B$ ,  $B_i$ , and  $B_j$  were, respectively, the total stand basal area, the basal area of trees of species  $i$ , and the basal area of trees in diameter class  $j$ ; and  $ns$  and  $nd$  represented total number of tree species and diameter classes. With 7 species and 20 2.54-cm (1-in) diameter classes being studied,  $ns = 7$  and the theoretical range of  $H_s$  was between 0 and  $\ln(7) = 1.95$ , whereas  $nd = 20$  and the theoretical range of  $H_d$  was between 0 and  $\ln(20) = 3.00$ . There was little correlation between  $H_s$  and  $H_d$ , so tree species diversity and tree size-class diversity represent two distinct measures of forest diversity.

To take into account the expected changes in tree species and tree size-class diversities ( $H_s$  and  $H_d$  respectively) from their current ecological state to a possible future ecological state, we separately applied the law of population growth (Lotka 1925) to  $H_s$  and  $H_d$  on each plot:

$$\begin{aligned} H(t) \\ = \frac{KH_0 e^{rt}}{K + H_0(e^{rt} - 1)} \end{aligned} \quad (3.3)$$

where  $K$  is equal to the theoretical maximum for either  $H_s$  or  $H_d$  as derived from this data set and presented above,  $H_0$  is equal to the value of  $H_s$  or  $H_d$  in the 1<sup>st</sup> inventory,  $r$  is the annual change of  $H_s$  or  $H_d$  calculated from the 1<sup>st</sup> and 2<sup>nd</sup> PSP inventories from the WAIN and CAFI datasets, and  $t$  is the time between the current and future state, in this case 40 years. These calculations incorporate the changes in basal area for the different tree species and size-classes due to mortality, recruitment, and growth that were observed between the two inventory periods.

### 3.3.2 ENVIRONMENTAL FACTORS

Our predictor dataset consisted of 28 variables, including climate, soil, distance from various features, and topographic variables, as well as the site coordinates (Table 3.1). We did not, however, include any variables which may potentially change over time (i.e., permafrost, vegetation cover, normalized difference vegetation index (NDVI), etc.) leading to unknown values. These data were derived from two different open access databases. The climate data were from the Scenarios Network for Alaska Planning database (SNAP; <http://www.snap.uaf.edu>), while the soil, distance, and topographical variables were created from data within the database maintained by USFS/State and Private Forestry and USGS Alaska Geographic Science Office (Alaska Geospatial Data Clearinghouse (AGDC); <http://agdc.usgs.gov/data>).

The current climate conditions came from the SNAP dataset which were derived from Climate Research Unit (CRU) data which has been shown to perform well in Alaska (Walsh et al. 2008). The spatial resolution of these monthly and annual temperature data is 2km, and data were averaged over the years 1901-2009, while the monthly and annual precipitation data were averaged from 1901-2006. For the future climate projections, the SNAP data were derived from general-circulation models (GCMs) used by the Intergovernmental Panel on Climate Change (IPCC) (Walsh et al. 2008). These data contain monthly and annual temperature and precipitation from projections covering the years 1980-2099 using five different climate models that provided the highest level of accuracy for Alaska (Walsh et al. 2008). For our analysis, we used the midrange (IPCC A1B scenario) greenhouse gas emission scenario for the year 2050, which assumes continued development, but aggressive efforts to reduce fossil fuel emissions. For the Euclidean distance variables we used Spatial Analysis surface analysis tools within ArcGIS 10.0 (ESRI 2011) using data within AGDC. The topographic variables were derived from 300m digital elevation models from AGDC using Spatial Analysis surface analysis tool and the topographic position index (TPI) extension for ArcGIS (Jenness 2006). The predictor variables with a spatial resolution greater than 1km<sup>2</sup> underwent either nearest neighbor resampling, if the data were categorical, or bilinear interpolation

resampling, if the data were continuous. Variables that had a native spatial resolution smaller than  $1\text{km}^2$  were rescaled. The predictor dataset was lastly constructed by overlaying the individual datasets in ArcGIS 10.0, then at each PSP location the environmental variables were extracted, as per Ohse et al. (2009). This resulted in a table with  $H_s$  and  $H_d$  values as the response variables and the environmental variables as predictors.

### 3.3.3 THE CALIBRATION AND VALIDATION DATASETS

For each of the 704 sites (PSPs) within our study region we had information about the diversity measures ( $H_s$  and  $H_d$ ) and the 26 environmental factors plus the X and Y coordinates (Table 3.1). This dataset was randomly split into a calibration dataset (Cal,  $n = 528$ ; 75% of the plots) and a validation dataset (Val,  $n = 176$ ; 25% of the plots). The correlations between the diversity measures and the environmental factors were modeled using the calibration dataset and the qualities of the predictions were assessed using the validation dataset. Neither the validation nor the calibration datasets accounted for landscape-level mortality due to wildfire or insect outbreaks.

### 3.3.4 STATISTICAL METHODS

Our modeling approach involved determining the association among  $H_s$ ,  $H_d$  and the environmental predictors at each of the PSP locations. The environmental predictors were first tested using hierarchical cluster analysis in the Hmisc package (Harrell 2009) for the R system (R Development Core Team 2010), using the squared Spearman correlations for assessing collinearity. The presence of multicollinearity typically produces spurious results when using regression models (Mac Nally 2002), but the competition between similar predictor variables can be reduced through random selection (Siroky 2009). To account for known complex ecological and environmental interactions among variables and limit the problems associated with multicollinearity (Cutler et al. 2007; De'ath and Fabricius 2000; Prasad et al. 2006; Siroky 2009), we used Random Forest Analysis (RFA; Breiman 2001) in the randomForest package (Liaw and Wiener



2002) for the R system. RFA can be used to estimate and rank the importance of the predictors in the relationships between the predictor and response variables.

The most relevant environmental predictors for  $H_s$  and  $H_d$  were selected using an algorithm developed for RFA for the R system called Boruta (Kursa and Rudnicki 2010). This algorithm adds several random permutations of each original variable by running the Random Forest algorithm multiple times (*maxRuns*), in this case 500 times. The statistical significance of each variable is then determined by comparing the original variable to the random permutation based on the Z-score with an upper 95% confidence limit ( $|Z| \leq 1.65$ ) (Kursa and Rudnicki 2010).

After conducting variable selection using the Boruta algorithm (Table 3.2), the RFA as applied to these data used 500 bootstrap samples (*ntree*) each containing two-thirds of the Cal data. The observations that were not included in each bootstrap sample are called out-of-bag (OOB) observations. For each bootstrap sample, an un-pruned regression tree was grown containing one-third of the predictor variables (*mtry*), which in this case were 6 at most, which were randomly selected and used for binary partitioning. The average of all trees was then used to predict OOB observations, which, in turn, allowed for cross validation (Breiman 2001) and the evaluation of the overall error of the RFA models. The importance values for each predictor was also calculated by investigating the percent increase in mean squared error (MSE) when OOB data for each variable were permuted while all others were kept constant (Breiman 2001; Cutler et al. 2007; Liaw and Wiener 2002). An external validation of the predictive capabilities of the RFA models was also conducted on an independent dataset (the Val data set with 176 measures of diversity and predictor variables) that was not used in the calibration of the RFA models. Pearson's product-moment correlation coefficients ( $r$ ) and the root mean square error (*RMSE*) were used as measures of model performance. In addition, partial dependence plots were developed to provide a way to visualize the marginal effects of the predictor variables in the RFA estimates of tree size-class and tree species diversities.

Considering that spatial autocorrelation affects statistical model predictions (Legendre 1993), we tested for spatial autocorrelation as well as for large-scale spatial

patterns within the residuals of the final RFA models. We assumed that plots at distant locations will affect each other less than plots that are close to one another (Cressie 1993). We therefore, applied a spatial weight of inverse distance. Given the neighborhood structure, we then evaluated the residuals of the RFA models using Moran's I and Geary's C test statistics (Sokal and Oden 1978) with the spdep (Bivand et al. 2007) package for the R system.

### 3.3.5 PREDICTIVE MAPS

Once calibrated and validated, the final RFA models (Table 3.2) were applied to the entire boreal forest region of Alaska at a 1km<sup>2</sup> resolution to obtain an estimate of the potential diversity values for the current year (2010) and forecast to 2050 for both a null climate change scenario (which allows for successional change in the absence of climate change) and the IPCC A1B "mid-range" scenario that includes the effects of both succession and climate change.

## 3.4 RESULTS

### 3.4.1 PREDICTOR VARIABLE CORRELATIONS

The hierarchical cluster analysis revealed strong correlations among several of the temperature and precipitation variables (Figure 3.2). These strong correlations among the predictor variables further support the use of the Random Forest Analysis (RFA; Breiman 2001) rather than using linear and parsimonious analysis types.

### 3.4.2 VARIABLE SELECTION AND IMPORTANCE

The variable selection process, as obtained from Boruta algorithm, yielded slightly different results for the variables that best predicted tree size-class diversity ( $H_d$ ) tree species diversity ( $H_s$ ) (Table 3.2). Each of the diversity measures shared numerous predictor variables. However, the variables of slope, solar insolation, and mean June temperature were only found to be significant predictors of  $H_d$  while, the variables of

distance to roadway, mean September temperature, and mean growing season temperature only predicted  $H_s$ . The final combinations of variables used in the RFA were the ones deemed important through this variable selection process (Table 3.2).

Table 3.2 also shows the ranking of predictors by their importance as determined by percent increase in mean standard error (%IncMSE) for both tree size-class diversity ( $H_d$ ) and tree species diversity ( $H_s$ ). For  $H_d$ , the first six variables (distance to navigable waterway, distance to community, elevation, Y-coordinate, aspect, and site productivity) each contributed over 15%IncMSE. The remaining 15 variables ranged in their importance from 13.92% down to 6.33%. For  $H_s$ , the first 11 variables (Y-coordinate, aspect, X-coordinate, distance to community, elevation, slope, site productivity, winter precipitation, mean annual temperature, June precipitation, and annual precipitation) each contributed over 15%IncMSE with the remaining variable importance ranging from 14.93% down to 9.77%. The overall influence of the predictor variables was greater for  $H_s$  compared to  $H_d$  as observed through the larger overall values of the predictor variables percent increase in mean standard error.

### 3.4.3 SPATIAL DEPENDENCY OF TREE SIZE-CLASS AND SPECIES DIVERSITY

Both tree size-class diversity ( $H_d$ ) and tree species diversity ( $H_s$ ) are strongly spatially autocorrelated (Table 3.3). By incorporating numerous environmental predictor variables in the final RFA models, the spatial autocorrelation present in  $H_d$  and  $H_s$  was effectively eliminated as evident by the lack of autocorrelation present in the residuals of the final models (Table 3.3).

### 3.4.4 RANDOM FOREST ANALYSIS MODEL ASSESSMENT FOR TREE SIZE-CLASS AND TREE SPECIES DIVERSITY

The RFA models (Table 3.2) as applied to the calibration dataset (Cal) for predicting tree size-class diversity ( $H_d$ ) and tree species diversity ( $H_s$ ) explained a large portion of the variance in  $H_d$  and  $H_s$  (Figure 3.3:  $r_{cal} = 0.935$ ,  $P < 0.001$ ; and Figure 3.4:  $r_{cal} = 0.934$ ,  $P < 0.001$ , respectively). The RFA models also provided highly significant

predictions for  $H_d$  and  $H_s$  in the validation dataset (Figure 3.3:  $r_{\text{val}} = 0.420$ ,  $P < 0.001$ ; and Figure 3.4:  $r_{\text{val}} = 0.446$ ,  $P < 0.001$ , respectively). However, despite these high significance levels, the RFA models for  $H_d$  and  $H_s$  both tended to overestimate diversity values for low diversity sites and underestimate it in high diversity sites (Figures 3.3:  $RMSE_{\text{val}} = 0.385$ ; and Figure 3.4:  $RMSE_{\text{val}} = 0.314$ , respectively).

Partial plots representing the marginal effects of the 12 most influential variables included in the RFA models to predict  $H_d$  and  $H_s$  are shown in Figures 3.5 and 3.6, respectively. Several of the correlations among the prediction variables and  $H_d$  and  $H_s$  are nonlinear. Additionally, some of the variables exhibit thresholds (Figure 3.5, distance to navigable waterways (DTW); Figure 3.6, June precipitation). The climatic variables are of particular importance considering that these are predicted to change and thus affect the levels of  $H_d$  or  $H_s$  modeled to the year 2050.

The three distance variables of distance to navigable waterways (Dtw), distance to community (Dtc), and distance to roadway (Dtr) are all highly influential in predicting tree size-class diversity ( $H_d$ ) with each contributing over 13% IncMSE (Table 3.2; Figure 3.5). These distance variables can be thought of as measures of the human impact on  $H_d$ . Both Dtc and Dtr initially exhibit a strong increase in  $H_d$  with increasing distance, while Dtw exhibits a decreased level of  $H_d$  starting at  $\sim 15\text{km}$  from the waterway (Figure 3.5). These measures of human impact combined with the topographic variables and site coordinates account for the eight most influential parameters that primarily dictate the observed patterns of  $H_d$  within the sampled sites.

The spatial structure, i.e., Y-coordinate (N) and X-coordinate (E), have a pronounced effect on tree species diversity ( $H_s$ ) (Table 3.2; Figure 3.6). The strong increase in  $H_s$  with increasing N is, however, likely a result of the extremely low  $H_s$  values within the PSPs driven by past disturbance events that occurred on the Kenai Peninsula and the Copper River valley (Allen et al. 2006; Boucher and Mead 2006) and perhaps not a result of any climatic trends. The strong positive correlation of E with  $H_s$  exhibits a strong east-west gradient in continentality. Transformed aspect (As) also contributed a sizable portion of the explained variance in  $H_s$  (Table 3.2) and exhibits a

strong negative correlation with  $H_s$ , with -1 (southwest facing slopes) having the highest  $H_s$  values (Figure 3.6).

### 3.4.5 CURRENT PREDICTED DIVERSITY PATTERNS

#### 3.4.5.1 TREE SIZE-CLASS DIVERSITY

The current levels of tree size-class diversity ( $H_d$ ) as predicted by the RFA model varied greatly across Alaska (Figure 3.7).  $H_d$  values ranged from greater than 2.04 (Shannon's index) to less than 1.58 across broad areas within the study region. The highest levels of diversity occur in a large area in the north eastern portion of the State and, a narrow band running towards the south west. The high levels of  $H_d$  in these regions are likely influenced by the close proximity to the upper Yukon River and its tributaries made evident by the high importance of the distance to navigable waterway (DTW) variable. Additionally, localized high levels of diversity are also observed within some portions of the Matanuska-Susitna River valleys north of Anchorage and the Copper River valley near Glennallen. The lowest  $H_d$  values occur in the Tanana River valley running between Tok and Fairbanks as well as along the lower Yukon and Koyukuk River valleys north and west of McGrath. These regions, despite their close proximity to navigable waterways, are more strongly influenced by other factors, such as their relatively low elevation and close proximity to developed communities (DTC), which appear to counteract the influence of DTW. Other regions of low  $H_d$  are also observed surrounding Anchorage and west of Glennallen, which are influenced by a greater DTW.

#### 3.4.5.2 TREE SPECIES DIVERSITY

The current level of tree species diversity ( $H_s$ ), as predicted by the RFA model, also varies considerably across Alaska (Figure 3.8). The range of  $H_s$  across the study region is from a low of less than 0.30 (Shannon's diversity index) to greater than 0.75. The lowest  $H_s$  levels, as predicted from the RFA model, are located on the eastern side of the Kenai Peninsula and the southern portion of the Copper River valley south east of

Glennallen. The low levels of  $H_s$  for these two regions is most likely a result of past disturbance events (Allen et al. 2006; Boucher and Mead 2006). Other areas that express low levels of  $H_s$  include the confluence of the Yukon and Koyukuk rivers west of McGrath, which appears to be driven by the strong east-west gradient in continentality and, a small area due west of Tok, which has an aspect primarily facing north-east. The highest levels of diversity can be found mainly throughout the Tanana River valley between Tok and Fairbanks and in localized pockets dispersed across the interior. For these regions, the southwest-facing hillsides with moderate slopes near the Tanana River are the major factors yielding the high  $H_s$  values despite the proximity to numerous communities. Unlike what was observed for  $H_d$ , the distance variable of distance to community ( $D_{tc}$ ), exhibits a positive correlation with  $H_s$ , and is particularly pronounced in this model as indicated by the apparent circles found throughout the study region (Figure 3.8).

### 3.4.6 PREDICTED FUTURE DIVERSITY PATTERNS ASSUMING NO MAJOR DISTURBANCE EVENTS

#### 3.4.6.1 TREE SIZE-CLASS DIVERSITY

Under the two different scenarios; null climate change (Figure 3.9A) and the IPCC A1B (Figure 3.9B), for tree size-class diversity ( $H_d$ ) for the year 2050, the vast majority of the study region is projected to experience a net positive change from its current level (Figure 3.7). For the null climate change model (Figure 3.9A), significant portions of the region are predicted to experience increased levels of  $H_d$  as a result of niche partitioning as the stands mature, however, within localized regions, within the extreme southeast, surrounding Fairbanks, and in scattered areas throughout the south-central region,  $H_d$  is projected to decrease from its current level (Figure 3.7). The projected decreases in  $H_d$  are likely the result of these stands reaching a more stable late successional state resulting in a restriction of niche breadth via increased competition for

resources. These projections do not account for fire, insects, and other disturbances that would produce early successional stands.

For the IPCC A1B scenario (Figure 3.9B), the level of increase in  $H_d$  is overall less than is predicted for the null model showing that a changing climate will slow forest succession, at least in terms of differentiation of  $H_d$ . The largest increases appear to be located primarily between the Yukon and Koyukuk rivers north and west of McGrath and in portions of the Tanana River valley east of Fairbanks. Additionally, fairly high increases are predicted for the Yukon Flats area located in the north-eastern portion of the region and localized areas within south-central and the Kenai Peninsula, indicating that the combined predicted temperature and precipitation changes for these areas are still largely favorable for increases in  $H_d$  to occur. However, significant reductions are predicted for areas south of Fairbanks and in the upper Copper River valley surrounding Glennallen, suggesting that projected climate change for these areas might cause a decrease in spatial resource partitioning or increased mortality of larger trees.

#### 3.4.6.2 TREE SPECIES DIVERSITY

Projected tree species diversity ( $H_s$ ) for the year 2050 under a null climate scenario (Figure 10A) and the IPCC 1AB scenario (Figure 3.10B) produce strikingly different results. For the null climate scenario (Figure 3.10A), the Forty Mile and Yukon Flats area, located in the north-eastern portion of the study region, is predicted to experience a large increase in species diversity. Additionally, scattered areas in the central and western interior and the southern Copper River valley surrounding Glennallen and on the southern tip of the Kenai Peninsula are also predicted to experience increases in tree species diversity from their current level. These increases in tree species diversity are primarily due to ingrowth of new species into the canopy as a result of gap phase successional processes in the absence of large scale disturbance. Large decreases in diversity are predicted for the null climate change scenario for tree species diversity for the Tanana River valley between Tok and Fairbanks as well as the area between the Yukon and Kuskokwim rivers in the south-west near McGrath. A marked reduction is

also predicted for the Matanuska-Susitna valley north of Anchorage. These predicted decreases are likely attributable to tree stands reaching a more stable late successional state resulting in decreased species richness.

For the IPCC A1B scenario (Figure 3.10B), the change in tree species diversity from its current level is as a whole far more negative than the null climate scenario (Figure 3.10A). The largest predicted decreases in tree species diversity are expected to occur in the Tanana River valley from Tok to Fairbanks and in the western interior between the Koyukuk and the Yukon River valleys. Additionally, small isolated areas in the Matanuska-Susitna valley and in the upper Copper River basin surrounding Glennallen. These predicted decreases in  $H_s$  are likely from synergistic correlations between the predicted climate change for these regions and local successional processes affecting rates of recruitment, growth, and mortality in the absence of major disturbance. Similar to the null climate scenario, the largest expected increase in  $H_s$  is predicted to occur in the Forty Mile and Yukon Flats area located in the north-east portion of the study region. The increases in  $H_s$  in these regions are due to the interactions of the predicted changes in climate with the successional processes yielding higher species diversity. The higher species diversity values on the Kenai Peninsula and the southernmost portion of the Copper River valley southeast of Glennallen are primarily due to the recovery from past disturbance events and the subsequent ingrowth of new species into the canopy (Allen et al. 2006; Boucher and Mead 2006).

### 3.5 DISCUSSION

The data from the forest inventories (CAFI and WAIN datasets) were critical to our analysis in determining the status of tree species and size-class diversities under both current and projected future conditions. While species range maps do exist for the study region (see Ohse et al. 2009; Ruefenacht et al. 2008; Viereck and Little 2007), they lack the information necessary to estimate regional patterns of diversity. The data contained within the CAFI and WAIN datasets comprise the largest collection of field data on forest dynamics in Alaska (Malone et al. 2009; Rees, personal communication). However, these



data are not uniformly dispersed across the study region (Figure 3.1). To this end, Random Forest Analysis (RFA; Breiman 2001) is ideally suited to studies such as this where other spatial prediction techniques (i.e., kriging) would potentially produce a single mean value for large portions of the study region.

The application of the RFA model to predict tree species and tree size-class diversities is particularly useful when there are complex ecological interactions among predictors and response variables and in the presence of highly correlated predictor variables (as evident within our data set, Figure 3.2). RFA models are also advantageous because they do not require the assumption of normality of the modeling variables, and they can be applied to non-linear relationships. Additionally, RFA models are highly useful due to their ability to make fairly accurate predictions for areas where observed data are lacking (Magness et al. 2008; Parmentier et al. 2011), such as the vast majority of our study region. In addition, RFA presents a time advantage and convenience, e.g., when compared to GLMs and parsimonious analysis (Huettmann et al. 2011).

We chose not to include variables in our model that have previously been used in determining landscape richness from remotely sensed data (i.e. normalized difference vegetation index (NDVI); Feilhauer and Schmidtlein 2009; Parmentier et al. 2011) because any potential future values would be based on predictions that, as far as we know, have not occurred for this study region. While this decision may have lowered our ability to make the most accurate predictions for the current levels of tree size-class and species diversities ( $H_d$  and  $H_s$ , respectively), it provides a more valid basis for predicting future conditions. Additionally, this choice to exclude some variables may have led to our overestimates for low levels and underestimates for high levels for both tree size-class ( $H_d$ ) and species ( $H_s$ ) diversities (Figures 3.3 and 3.4 respectively). This trend in over- and underestimation has been observed in other studies that employed RFA for making predictions (Baccini et al. 2004; Baccini et al. 2008; Vincenzi et al. 2011). This trend is likely the result of the predictions being based on the average values within the terminal nodes within tree-based models which occur when the splitting procedure stops (Breiman

2001) or the learning rate of the Random Forest algorithm was too quick (Friedman 2001).

The predictions for the current tree size-class ( $H_d$ ) diversity (Figure 3.7) at the  $1\text{km}^2$  scale, as presented in this study, are strongly influenced by the distance variables: distance to navigable water, distance to roads, and distance to communities. This shows that these variables may represent the human impact from historical logging in interior Alaska (Wurtz and Gasbarro 1996; Wurtz et al. 2006). This result suggests that the human impacts were better predictors than climatic or environmental variables, i.e., the various state factors (Major 1951), which have previously been shown to drive diversity patterns within the boreal forest (Chapin et al. 2006a; Hooper et al. 2005). The influence of some of the topography variables on both  $H_d$  and  $H_s$  were also particularly pronounced (Table 3.2). It further confirms that these variables are major determinants of vegetation distribution within interior Alaska (Chapin et al. 2006b). The integrated climate proxies of Y-coordinate, X-coordinate, and elevation were also found to be of sizable importance in predicting the patterns of  $H_d$  and  $H_s$  which, at the scale of  $1\text{km}^2$  and greater, have previously been found to significantly explain patterns of diversity (Heikkinen 1996; Heikkinen et al. 2004; Kallimanis et al. 2007).

Through succession, the diversity of vascular plants within the Alaskan boreal forests tends to decline (Chapin et al. 2006b). Within this forest, succession typically leads to a change from deciduous dominance to conifer dominance due to the recruitment of conifers into the canopy (Bergeron et al. 2002; Chapin et al. 2006b; Harper et al. 2006). The areas within our study of reduced tree size-class ( $H_d$ ) and tree species ( $H_s$ ) diversity under a null climate change scenario is likely due to succession. However, post-disturbance recruitment in the boreal forest depends on both the severity and extent of a disturbance event (Johnstone and Chapin 2006). For example, wildfire severity directly affects stand structure and composition because post-fire recruits generally dominate the canopy (Gutsell and Johnson 2002) for years following the disturbance. Additionally, sites that have experienced high-severity wildfire may also exhibit increased abundance and richness of understory vegetation (Bernhardt et al. 2011), which may be driving the

gains of  $H_d$  and  $H_s$  in our study (Figures 3.9a and 3.10a). The changes in forest succession, as presented here (Figures 3.9a and 3.10a), only account for the various levels of disturbances that were present within the permanent sampling plots of the CAFI and WAIN datasets and do not incorporate projected future disturbance events.

Previous studies have found strong correlations between the patterns of plant richness and climate (Francis and Currie 2003; Hawkins et al. 2003; Hawkins et al. 2007; Kreft and Jetz 2007; O'Brien 1998). However, these studies all investigated plant richness at very large scales ( $> 20\text{km}^2$ ), whereas, in this analysis, we investigated the patterns of diversity at a  $1\text{km}^2$  grid. We found that most of the variation in tree size-class reflected distance variables (human impact), whereas tree species diversity was explained more strongly by local topographic variables (Table 3.2). Of the climatic variables, June precipitation (P\_06) was the most important in predicting tree size-class diversity ( $H_d$ ), whereas winter precipitation (P\_W) was the most important for tree species diversity ( $H_s$ ) (Table 3.2). These moisture variables have previously been linked as primary limiting factors to forest growth within the Alaskan boreal forest (Yarie and Van Cleve 2010) and diversity in general (Hawkins et al. 2003; Kreft and Jetz 2007). The predicted changes in both temperature and precipitation as forecast in the IPCC A1B “mid-range” model scenario vary spatially across the Alaskan boreal forest (Walsh et al. 2008). The effects of a changing climate impact the patterns of diversity differently (Figures 3.9b and 3.10b) depending on location due to complex interactions of the state factors (Chapin et al. 2004; Chapin et al. 2006b; Yarie and Van Cleve 2010).

The predicted changes in forest diversity, as modeled here (Figures 3.9 and 3.10), are based on the assumption that forest stand conditions will respond in the future in ways that are similar to past variation in climate without barriers to species migration. These assumptions may be reasonable in the short term but may not be the case for longer term, considering that there are multiple successional pathways in the boreal forest (Fastie 1995; Frelich and Reich 1995; Hollingsworth et al. 2010; Johnstone et al. 2011; Taylor and Chen 2011) that are largely determined by the dominant species composition and the time since the last fire event (Johnstone et al. 2010; Taylor and Chen 2011).

Long-term predictions are particularly difficult to validate because feedbacks that come into play over long time scales could give rise to dynamics that are not well represented in the short term (i.e. species migration, fire return interval; Soja et al. 2007; Yarie 1981). Perhaps even more important, the model presented in this paper allows stands to mature as a result of within-stand growth and mortality, but it does not incorporate stand-replacing disturbances such as fire and insect outbreaks which would increase the relative frequency of young stands, which have low size-class diversity.

Biodiversity is also spatially driven (Roberts and Gilliam 1995), so it is crucial to account for spatial autocorrelation when conducting landscape-level analysis (Bivand et al. 2008; Sokal and Oden 1978; Wagner and Fortin 2005). By not taking spatial autocorrelation into account, previous studies of diversity may have generated overestimates of the effects of autocorrelated explanatory variables (Legendre 1993; Lichstein et al. 2002). For example, Liang et al. (2011) incorporated “geographical dependence” (Miller et al. 2007), which may have explained the broad level trends in the data but did not include the spatial dependence of the data. The method proposed in this study incorporates the spatial dependence and could act as a model to assist forest managers and researchers interested in these dynamics.

Predictive mapping is a powerful tool for landscape-level planning and analysis (Franklin 1995). Improved understanding of the current and potential future landscape-level patterns of biodiversity may assist land management agencies in their decision making processes in regards to sustainability forestry activities (Ogden and Innes 2009). Our proposed predictive mapping of tree species and tree size-class diversity at the 1-km<sup>2</sup> scale would best be used for broad landscape-level planning. We believe that these models could be improved by adding more ground-sampled data, open access data sources, and a stronger collaboration with forest practitioners to assess if these models meet their particular needs, interests, or applications.

### 3.6 REFERENCES

Allen, J.L., S. Wesser, C.J. Markon, and K.C. Winterberger. 2006. Stand and landscape level effects of a major outbreak of spruce beetles on forest vegetation in the Copper River Basin, Alaska. *For. Ecol. Manage.* 227(3):257-266.

Austin, M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol. Model.* 157(2-3):101-118.

Baccini, A., M. Friedl, C. Woodcock, and R. Warbington. 2004. Forest biomass estimation over regional scales using multisource data. *Geophysical Research Letters* 31(10):L10501.

Baccini, A., N. Laporte, S.J. Goetz, M. Sun, and H. Dong. 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters* 3(4):045011.

Bergeron, Y., A. Leduc, B.D. Harvey, and S. Gauthier. 2002. Natural fire regime: A guide for sustainable management of the Canadian boreal forest. *Silva Fennica* 36(1):81-95.

Bernhardt, E.L., T.N. Hollingsworth, and F.S. Chapin. 2011. Fire severity mediates climate-driven shifts in understory community composition of black spruce stands of interior Alaska. *J. Veg. Sci.* 22(1):32-44.

Bivand, R.S., L. Anselin, O. Berke, A. Bernat, M. Carvalho, Y. Chun, C.F. Dormann, S. Dray, R. Halbersma, and N. Lewin-Koh. 2007. *spdep: Spatial dependence: weighting schemes, statistics and models*. R package version 0.4-9.

Bivand, R.S., E.J. Pebesma, and V. Gómez-Rubio. 2008. *Applied spatial data analysis with R*. Springer Verlag, New York, New York, USA.

Boucher, T.V., and B.R. Mead. 2006. Vegetation change and forest regeneration on the Kenai Peninsula, Alaska following a spruce beetle outbreak, 1987–2000. *For. Ecol. Manage.* 227(3):233-246.

Breiman, L. 2001. Random forests. *Mach. Learn.* 45(1):5-32.

Bunnell, F.L., and D.J. Huggard. 1999. Biodiversity across spatial and temporal scales: problems and opportunities. *For. Ecol. Manage.* 115(2-3):113-126.

Buongiorno, J., S. Dahir, H.C. Lu, and C.R. Lin. 1994. Tree Size Diversity and Economic Returns in Uneven-Aged Forest Stands. *For. Sci.* 40(1):83-103.

Chapin, F.S., 3rd, T.V. Callaghan, Y. Bergeron, M. Fukuda, J.F. Johnstone, G. Juday, and S.A. Zimov. 2004. Global change and the boreal forest: thresholds, shifting states or gradual change? *Ambio* 33(6):361-365.

Chapin, F.S., 3rd, A.L. Lovcraft, E.S. Zavaleta, J. Nelson, M.D. Robards, G.P. Kofinas, S.F. Trainor, G.D. Peterson, H.P. Huntington, and R.L. Naylor. 2006a. Policy strategies to address sustainability of Alaskan boreal forests in response to a directionally changing climate. *Proc Natl Acad Sci U S A* 103(45):16637-16643.

Chapin, F.S., III, T. Hollingsworth, D.F. Murray, L.A. Viereck, and M.D. Walker. 2006b. Floristic diversity and vegetation distribution in the Alaskan boreal forest. P. 81-99 in *Alaska's changing boreal forest*, Chapin, F.S., III, M. Oswood, K. Van Cleve, L.A. Viereck, and D. Verbyla (eds.). Oxford University Press, New York.

Chapin, F.S., M.W. Oswood, K. Van Cleve, L.A. Viereck, M.C. Chapin, and D.L. Verbyla (eds.). 2006c. *Alaska's changing boreal forest*. Oxford University Press, New York, New York, USA. 368 p.

Condit, R., P. Ashton, S. Bunyavejchewin, H.S. Dattaraja, S. Davies, S. Esufali, C. Ewango, R. Foster, I.A. Gunatilleke, C.V. Gunatilleke, P. Hall, K.E. Harms, T. Hart, C.

Hernandez, S. Hubbell, A. Itoh, S. Kiratiprayoon, J. Lafrankie, S.L. de Lao, J.R. Makana, M.N. Noor, A.R. Kassim, S. Russo, R. Sukumar, C. Samper, H.S. Suresh, S. Tan, S. Thomas, R. Valencia, M. Vallejo, G. Villa, and T. Zillio. 2006. The importance of demographic niches to tree diversity. *Science* 313(5783):98-101.

Craig, E., and F. Huettmann. 2008. Using “blackbox” algorithms such as TreeNet and Random Forests for data-mining and for finding meaningful patterns, relationships and outliers in complex ecological data: an overview, an example using golden eagle satellite data and an outlook for a promising future. in *Intelligent Data Analysis: Developing New Methodologies through Pattern Discovery and Recovery*, Wang, H.-f. (ed.). IGI Global, Hershey, PA, USA.

Crawley, M.J., and J.E. Harral. 2001. Scale dependence in plant biodiversity. *Science* 291(5505):864-868.

Cressie, N.A.C. 1993. *Statistics for spatial data*. John Wiley & Sons, New York, New York, USA.

Cushman, S., and F. Huettmann (eds.). 2010. *Spatial complexity, informatics, and wildlife conservation*. Springer, Tokyo, Japan. 464 p.

Cushman, S.A., and K.S. McKelvey. 2009. Data on distribution and abundance: monitoring for research and management. P. 111-129 in *Spatial complexity, informatics, and wildlife conservation*, Cushman, S.A., and F. Huettmann (eds.). Springer Tokyo, Japan.

Cutler, D.R., T.C. Edwards, Jr., K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J.J. Lawler. 2007. Random forests for classification in ecology. *Ecology* 88(11):2783-2792.

De'ath, G., and K.E. Fabricius. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology* 81(11):3178-3192.

- Drew, C.A., Y.F. Wiersma, and F. Huettmann (eds.). 2010. *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications*. Springer Verlag, New York, New York, USA. 314 p.
- ESRI. 2011. *ArcGIS Desktop: Release 10*. Environmental Systems Research Institute, Redlands, CA.
- Fassnacht, K.S., W.B. Cohen, and T.A. Spies. 2006. Key issues in making and using satellite-based maps in ecology: A primer. *For. Ecol. Manage.* 222(1-3):167-181.
- Fastie, C.L. 1995. Causes and Ecosystem Consequences of Multiple Pathways of Primary Succession at Glacier Bay, Alaska. *Ecology* 76(6):1899-1916.
- Feilhauer, H., and S. Schmidtlein. 2009. Mapping continuous fields of forest alpha and beta diversity. *Applied Vegetation Science* 12(4):429-439.
- Ferrier, S., G. Watson, J. Pearce, and M. Drielsma. 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. *Biodivers. Conserv.* 11(12):2275-2307.
- Francis, A.P., and D.J. Currie. 2003. A globally consistent richness-climate relationship for angiosperms. *Am Nat* 161(4):523-536.
- Franklin, J. 1995. Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. *Prog. Phys. Geogr.* 19(4):474-499.
- Franklin, J.F. 1988. Structural and functional diversity in temperate forests. P. 166–175 in *Biodiversity*, Wilson, E.O. (ed.). National Academy Press, Washington, D.C., USA.
- Frelich, L.E., and P.B. Reich. 1995. Spatial Patterns and Succession in a Minnesota Southern-Boreal Forest. *Ecol. Monogr.* 65(3):325-346.



- Friedman, J.H. 2001. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* 29(5):1189-1232.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8(9):993-1009.
- Gutsell, S.L., and E.A. Johnson. 2002. Accurately ageing trees and examining their height-growth rates: implications for interpreting forest dynamics. *J. Ecol.* 90(1):153-166.
- Harper, K.A., Y. Bergeron, P. Drapeau, S. Gauthier, and L. De Grandpre. 2006. Changes in spatial pattern of trees and snags during structural development in *Picea mariana* boreal forests. *J. Veg. Sci.* 17(5):625-636.
- Harrell, F.E., Jr. 2009. Hmisc: Harrell Miscellaneous. R package version 3.7-0.
- Hawkins, B.A., R. Field, H.V. Cornell, D.J. Currie, J.F. Guegan, D.M. Kaufman, J.T. Kerr, G.G. Mittelbach, T. Oberdorff, E.M. O'Brien, E.E. Porter, and J.R.G. Turner. 2003. Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84(12):3105-3117.
- Hawkins, B.A., D. Montoya, M.A. Rodriguez, M.A. Olalla-Tarraga, and M.A. Zavala. 2007. Global models for predicting woody plant richness from climate: comment. *Ecology* 88(1):255-259; discussion 259-262.
- Heikkinen, R.K. 1996. Predicting patterns of vascular plant species richness with composite variables: A meso-scale study in Finnish Lapland. *Vegetatio* 126(2):151-165.
- Heikkinen, R.K., M. Luoto, R. Virkkala, and K. Rainio. 2004. Effects of habitat cover, landscape structure and spatial variables on the abundance of birds in an agricultural-forest mosaic. *J. Appl. Ecol.* 41(5):824-835.

Hellmann, J.J., and G.W. Fowler. 1999. Bias, precision, and accuracy of four measures of species richness. *Ecol. Appl.* 9(3):824-834.

Hollingsworth, T.N., A.H. Lloyd, D.R. Noss, R.W. Ruess, B.A. Charlton, and K. Kielland. 2010. Twenty-five years of vegetation change along a putative successional chronosequence on the Tanana River, Alaska. This article is one of a selection of papers from *The Dynamics of Change in Alaska's Boreal Forests: Resilience and Vulnerability in Response to Climate Warming*. *Canadian Journal of Forest Research* 40(7):1273-1287.

Hooper, D.U., F.S. Chapin, J.J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J.H. Lawton, D.M. Lodge, M. Loreau, S. Naeem, B. Schmid, H. Setälä, A.J. Symstad, J. Vandermeer, and D.A. Wardle. 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecol. Monogr.* 75(1):3-35.

Hubbell, S.P. 2001. *The unified neutral theory of biodiversity and biogeography*. Princeton University Press, Princeton, N.J., USA.

Huettemann, F., Y. Artukhin, O. Gilg, and G. Humphries. 2011. Predictions of 27 Arctic pelagic seabird distributions using public environmental variables, assessed with colony data: a first digital IPY and GBIF open access synthesis platform. *Marine Biodiversity* 41(1):141-179.

Iverson, L.R., and A.M. Prasad. 2001. Potential changes in tree species richness and forest community types following climate change. *Ecosystems* 4(3):186-199.

Jenness, J. 2006. Topographic position index (tpi\_jen.avx) extension for Arcview 3.x, v.1.3a. <http://www.jennessent.com/arcview/tpi.htm>, Jenness Enterprises [EB/OL].

Johnstone, J., and F. Chapin. 2006. Effects of soil burn severity on post-fire tree recruitment in boreal forest. *Ecosystems* 9(1):14-31.

Johnstone, J.F., T.N. Hollingsworth, F.S. Chapin, and M.C. Mack. 2010. Changes in fire regime break the legacy lock on successional trajectories in Alaskan boreal forest. *Glob. Change Biol.* 16(4):1281-1295.

Johnstone, J.F., T.S. Rupp, M. Olson, and D. Verbyla. 2011. Modeling impacts of fire severity on successional trajectories and future fire behavior in Alaskan boreal forests. *Landsc. Ecol.* 26(4):487-500.

Jost, L. 2006. Entropy and diversity. *Oikos* 113(2):363-375.

Juday, G.P., V. Barber, P. Duffy, H. Linderholm, S. Rupp, S. Sparrow, E. Vaganov, J. Yarie, and a. al. 2005. Forest, land management, agriculture. P. 781-862 in *Arctic climate impact assessment*, Council, A. (ed.). Cambridge University Press, Cambridge, U.K.

Kallimanis, A.S., V. Ragia, S.P. Sgardelis, and J.D. Pantis. 2007. Using regression trees to predict alpha diversity based upon geographical and habitat characteristics. *Biodivers. Conserv.* 16(13):3863-3876.

Kasischke, E.S., D.L. Verbyla, T.S. Rupp, A.D. McGuire, K.A. Murphy, R. Jandt, J.L. Barnes, E.E. Hoy, P.A. Duffy, M. Calef, and M.R. Turetsky. 2010. Alaska's changing fire regime — implications for the vulnerability of its boreal forests This article is one of a selection of papers from *The Dynamics of Change in Alaska's Boreal Forests: Resilience and Vulnerability in Response to Climate Warming*. *Canadian Journal of Forest Research* 40(7):1313-1324.

Kneeshaw, D.D., A. Leduc, P. Drapeau, S. Gauthier, D. Pare, R. Carignan, R. Doucet, L. Bouthillier, and C. Messier. 2000. Development of integrated ecological standards of sustainable forest management at an operational scale. *For. Chron.* 76(3):481-493.

Kreft, H., and W. Jetz. 2007. Global patterns and determinants of vascular plant diversity. *Proc Natl Acad Sci U S A* 104(14):5925-5930.

Kursa, M.B., and W.R. Rudnicki. 2010. Feature Selection with the Boruta Package. *Journal of Statistical Software* 36(11):1-13.

Legendre, P. 1993. Spatial Autocorrelation - Trouble or New Paradigm. *Ecology* 74(6):1659-1673.

Lei, X.D., W.F. Wang, and C.H. Peng. 2009. Relationships between stand growth and structural diversity in spruce-dominated forests in New Brunswick, Canada. *Can. J. For. Res.-Rev. Can. Rech. For.* 39(10):1835-1847.

Li, J., A.D. Heap, A. Potter, and J.J. Daniell. 2011. Application of machine learning methods to spatial interpolation of environmental variables. *Environ Modell Softw* 26(12):1647-1659.

Liang, J.J. 2010. Dynamics and management of Alaska boreal forest: An all-aged multi-species matrix growth model. *For. Ecol. Manage.* 260(4):491-501.

Liang, J.J., J. Buongiorno, and R.A. Monserud. 2006. Bootstrap simulation and response surface optimization of management regimes for Douglas-fir/western hemlock stands. *For. Sci.* 52(5):579-594.

Liang, J.J., J. Buongiorno, R.A. Monserud, E.L. Kruger, and M. Zhou. 2007. Effects of diversity of tree species and size on forest basal area growth, recruitment, and mortality. *For. Ecol. Manage.* 243(1):116-127.

Liang, J.J., and M. Zhou. 2010. A geospatial model of forest dynamics with controlled trend surface. *Ecol. Model.* 221(19):2339-2352.

Liang, J.J., M. Zhou, D.L. Verbyla, L.J. Zhang, A.L. Springsteen, and T. Malone. 2011. Mapping forest dynamics under climate change: A matrix model. *For. Ecol. Manage.* 262(12):2250-2262.

Liaw, A., and M. Wiener. 2002. Classification and Regression by randomForest. *R News* 2(3):18-22.

Lichstein, J.W., T.R. Simons, S.A. Shiner, and K.E. Franzreb. 2002. Spatial autocorrelation and autoregressive models in ecology. *Ecol. Monogr.* 72(3):445-463.

Lindenmayer, D.B., C.R. Margules, and D.B. Botkin. 2000. Indicators of biodiversity for ecologically sustainable forest management. *Conserv. Biol.* 14(4):941-950.

Lloyd, A.H., and C.L. Fastie. 2002. Spatial and temporal variability in the growth and climate response of treeline trees in Alaska. *Climatic Change* 52(4):481-509.

Lotka, A.J. 1925. *Elements of Physical Biology*. Williams and Wilkins, Baltimore, Md. USA.

Mac Nally, R. 2002. Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables. *Biodivers. Conserv.* 11(8):1397-1401.

Magness, D., F. Huettmann, and J. Morton. 2008. Using Random Forests to Provide Predicted Species Distribution Maps as a Metric for Ecological Inventory & Monitoring Programs P. 209-229 in *Applications of Computational Intelligence in Biology*, Smolinski, T., M. Milanova, and A.-E. Hassanien (eds.). Springer Berlin.

Magurran, A.E. 1988. *Ecological diversity and its measurement*. Princeton University Press, Princeton, N.J., USA.

Major, J. 1951. A functional, factorial approach to plant ecology. *For. Ecol. Manage.* 32:392-412.

Malone, T., J. Liang, and E.C. Packee. 2009. Cooperative Alaska Forest Inventory. Gen. Tech. Rep. PNW-GTR-785. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 42 p.

McGuire, A.D., R.W. Ruess, A. Lloyd, J. Yarie, J.S. Klein, and G.P. Juday. 2010. Vulnerability of white spruce tree growth in interior Alaska in response to climate variability: dendrochronological, demographic, and experimental perspectives. *Can. J. For. Res.-Rev. Can. Rech. For.* 40(7):1197-1209.

McRoberts, R.E., S. Winter, G. Chirici, E. Hauk, D.R. Pelz, W.K. Moser, and M.A. Hatfield. 2008. Large-scale spatial patterns of forest structural diversity. *Can. J. For. Res.-Rev. Can. Rech. For.* 38(3):429-438.

Miller, J., J. Franklin, and R. Aspinall. 2007. Incorporating spatial dependence in predictive vegetation models. *Ecol. Model.* 202(3-4):225-242.

Niemelä, J. 1999. Management in Relation to Disturbance in the Boreal Forest. *For. Ecol. Manage.* 115:127-134.

Nightingale, J.M., W.H. Fan, N.C. Coops, and R.H. Waring. 2008. Predicting tree diversity across the United States as a function of modeled gross primary production. *Ecol. Appl.* 18(1):93-103.

O'Brien, E.M. 1998. Water-energy dynamics, climate, and prediction of woody plant species richness: an interim general model. *J. Biogeogr.* 25(2):379-398.

O'Neill, R.V., C.T. Hunsaker, K.B. Jones, K.H. Riitters, J.D. Wickham, P.M. Schwartz, I.A. Goodman, B.L. Jackson, and W.S. Baillargeon. 1997. Monitoring Environmental Quality at the Landscape Scale. *Bioscience* 47(8):513-519.

- Ogden, A.E., and J.L. Innes. 2009. Application of Structured Decision Making to an Assessment of Climate Change Vulnerabilities and Adaptation Options for Sustainable Forest Management. *Ecology and Society* 14(1):11.
- Ohse, B., F. Huettmann, S.M. Ickert-Bond, and G.P. Juday. 2009. Modeling the distribution of white spruce (*Picea glauca*) for Alaska with high accuracy: an open access role-model for predicting tree species in last remaining wilderness areas. *Polar Biol.* 32(12):1717-1729.
- Parmentier, I., R.J. Harrigan, W. Buermann, E.T.A. Mitchard, S. Saatchi, Y. Malhi, F. Bongers, W.D. Hawthorne, M.E. Leal, S.L. Lewis, L. Nusbaumer, D. Sheil, M.S.M. Sosef, K. Affum-Baffoe, A. Bakayoko, G.B. Chuyong, C. Chatelain, J.A. Comiskey, G. Dauby, J.-L. Doucet, S. Fauset, L. Gautier, J.-F. Gillet, D. Kenfack, F.N. Kouamé, E.K. Kouassi, L.A. Kouka, M.P.E. Parren, K.S.H. Peh, J.M. Reitsma, B. Senterre, B. Sonké, T.C.H. Sunderland, M.D. Swaine, M.G.P. Tchouto, D. Thomas, J.L.C.H. Van Valkenburg, and O.J. Hardy. 2011. Predicting alpha diversity of African rain forests: models based on climate and satellite-derived data do not perform better than a purely spatial model. *J. Biogeogr.* 38(6):1164-1176.
- Prasad, A.M., L.R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* 9(2):181-199.
- Rands, M.R., W.M. Adams, L. Bennun, S.H. Butchart, A. Clements, D. Coomes, A. Entwistle, I. Hodge, V. Kapos, J.P. Scharlemann, W.J. Sutherland, and B. Vira. 2010. Biodiversity conservation: challenges beyond 2010. *Science* 329(5997):1298-1303.
- Roberts, M.R., and F.S. Gilliam. 1995. Patterns and Mechanisms of Plant Diversity in Forested Ecosystems - Implications for Forest Management. *Ecol. Appl.* 5(4):969-977.

- Ruefenacht, B., M.V. Finco, M.D. Nelson, R. Czaplewski, E.H. Helmer, J.A. Blackard, G.R. Holden, A.J. Lister, D. Salajanu, D. Weyermann, and K. Winterberger. 2008. Conterminous US and Alaska Forest Type Mapping Using Forest Inventory and Analysis Data. *Photogramm. Eng. Remote Sens.* 74(11):1379-1388.
- Serreze, M.C., J.E. Walsh, F.S. Chapin, T. Osterkamp, M. Dyurgerov, V. Romanovsky, W.C. Oechel, J. Morison, T. Zhang, and R.G. Barry. 2000. Observational evidence of recent change in the northern high-latitude environment. *Climatic Change* 46(1-2):159-207.
- Shannon, C.E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27:379-423.
- Siroky, D.S. 2009. Navigating Random Forests and related advances in algorithmic modeling. *Statistics Surveys* 3(0):147-163.
- Soja, A.J., N.M. Tchebakova, N.H.F. French, M.D. Flannigan, H.H. Shugart, B.J. Stocks, A.I. Sukhinin, E.I. Parfenova, F.S. Chapin, and P.W. Stackhouse. 2007. Climate-induced boreal forest change: Predictions versus current observations. P. 274-296. Elsevier Science Bv.
- Sokal, R., and N. Oden. 1978. Spatial autocorrelation in biology 1. Methodology. *Biological Journal of the Linnean Society* 10(2):199-228.
- Taylor, A.R., and H.Y.H. Chen. 2011. Multiple successional pathways of boreal forest stands in central Canada. *Ecography* 34(2):208-219.
- Team, R.D.C. 2010. R: A language and environment for statistical computing. in R Foundation for Statistical Computing, Vienna, Austria.



- Tilman, D., J. Knops, D. Wedin, P. Reich, M. Ritchie, and E. Siemann. 1997. The influence of functional diversity and composition on ecosystem processes. *Science* 277(5330):1300-1302.
- Van Cleve, K., C.T. Dyrness, L.A. Viereck, J. Fox, F.S. Chapin, and W. Oechel. 1983. Taiga ecosystems in interior Alaska. *Bioscience* 33(1):39-44.
- Varga, P., H.Y.H. Chen, and K. Klinka. 2005. Tree-size diversity between single- and mixed-species stands in three forest types in western Canada. *Can. J. For. Res.-Rev. Can. Rech. For.* 35(3):593-601.
- Viereck, L., and E. Little. 2007. Alaska trees and shrubs. University of Alaska Press, Fairbanks, Alaska, USA.
- Vincenzi, S., M. Zucchetta, P. Franzoi, M. Pellizzato, F. Pranovi, G.A. De Leo, and P. Torricelli. 2011. Application of a Random Forest algorithm to predict spatial distribution of the potential yield of *Ruditapes philippinarum* in the Venice lagoon, Italy. *Ecol. Model.* 222(8):1471-1478.
- Wagner, H.H., and M.J. Fortin. 2005. Spatial analysis of landscapes: Concepts and statistics. *Ecology* 86(8):1975-1987.
- Walsh, J.E., W.L. Chapman, V. Romanovsky, J.H. Christensen, and M. Stendel. 2008. Global Climate Model Performance over Alaska and Greenland. *J. Clim.* 21(23):6156-6174.
- Whittaker, R.J., K.J. Willis, and R. Field. 2001. Scale and species richness: towards a general, hierarchical theory of species diversity. *J. Biogeogr.* 28(4):453-470.
- Wiens, J.A. 1989. Spatial Scaling in Ecology. *Funct Ecol* 3(4):385-397.

- Wilmking, M., and G.P. Juday. 2005. Longitudinal variation of radial growth at Alaska's northern treeline - recent changes and possible scenarios for the 21st century. P. 282-300. Elsevier Science Bv.
- Wurtz, T.L., and A.F. Gasbarro. 1996. A brief history of wood use and forest management in Alaska. *For. Chron.* 72(1):47-50.
- Wurtz, T.L., R.A. Ott, and J.C. Maisch. 2006. Timber harvest in interior Alaska. P. 302-308 in *Alaska's Changing Boreal Forest*, Chapin, F.S., III, M. Oswood, K. Van Cleve, L. Viereck, and D. Verbyla (eds.). Oxford University Press, New York, New York.
- Yarie, J. 1981. Forest Fire Cycles and Life-Tables - a Case-Study from Interior Alaska. *Can. J. For. Res.-Rev. Can. Rech. For.* 11(3):554-562.
- Yarie, J., and K. Van Cleve. 2010. Long-term monitoring of climatic and nutritional affects on tree growth in interior Alaska. *Can. J. For. Res.-Rev. Can. Rech. For.* 40(7):1325-1335.
- Young, B., J.J. Liang, and F.S. Chapin. 2011. Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach. *For. Ecol. Manage.* 262(8):1608-1617.

Table 3.1: Definition of variables used in the analysis

Variable	Description	Unit	Reference
<b>Dependent variables</b>			
H <sub>d</sub>	Tree size-class basal area diversity	Shannon's index	Shannon (1948)
H <sub>s</sub>	Tree species basal area diversity	Shannon's index	Shannon (1948)
<b>Prediction variables</b>			
N	X- coordinate (Alaska Albers)	10 <sup>5</sup> m	Magness et al. (2008)
E	Y- coordinate (Alaska Albers)	10 <sup>5</sup> m	Magness et al. (2008)
T_01	Mean temperature January	(°C + 100)	Ohse et al. (2009)
T_05	Mean temperature May	(°C + 100)	Ohse et al. (2009)
T_06	Mean temperature June	(°C + 100)	Ohse et al. (2009)
T_07	Mean temperature July	(°C + 100)	Ohse et al. (2009)
T_08	Mean temperature August	(°C + 100)	Ohse et al. (2009)
T_09	Mean temperature September	(°C + 100)	Ohse et al. (2009)
T_G	Mean temperature growing season (May-September)	(°C + 100)	Ohse et al. (2009)
T_D	Mean temperature difference July-January	(°C + 100)	Ohse et al. (2009)
T_A	Mean annual temperature	(°C + 100)	Ohse et al. (2009)
P_05	Precipitation sum May	mm	Ohse et al. (2009)
P_06	Precipitation sum June	mm	Ohse et al. (2009)
P_07	Precipitation sum July	mm	Ohse et al. (2009)
P_08	Precipitation sum August	mm	Ohse et al. (2009)
P_09	Precipitation sum September	mm	Ohse et al. (2009)
P_G	Precipitation sum growing season (May-September)	mm	Ohse et al. (2009)
P_W	Precipitation sum winter (October-April)	mm	Yarie (2008)
P_A	Precipitation sum annual	mm	Ohse et al. (2009)
Solar	Potential maximum solar insolation	(kWH/m <sup>2</sup> )	Fu and Rich (1999)
Prod	Site productivity	unitless	Stage and Salas (2007)
Sl	Slope	percent	Stage and Salas (2007)
As	Transformed aspect	unitless	Beers et al. (1966)
El	Elevation	m	Magness et al. (2008)
Dtw	Distance to navigable waterway	km	Wurtz et al. (2006)
Dtc	Distance to community	km	Wurtz et al. (2006)
Dtr	Distance to roadway	km	Wurtz et al. (2006)
Soil	Soil type	class	Ohse et al. (2009)

**Table 3.2: Variable importance (ranking) in determining tree size-class diversity ( $H_d$ ) and tree species diversity ( $H_s$ ) using percent increase in mean standard error (%IncMSE) for ranking purposes. All variables listed in this table were determined to be significant based on the Z-score with an upper 95% confidence limit ( $|Z| \leq 1.65$ ) and were used in the development of the final random forest analysis models.**

Variables	Tree Size-Class Diversity		Tree Species Diversity	
	%IncMSE	rank	%IncMSE	rank
Dtw	18.47	1	14.47	14
Dtc	17.97	2	20.72	4
El	17.19	3	19.82	5
N	16.78	4	26.34	1
As	16.72	5	21.96	2
Prod	15.51	6	18.94	7
E	13.92	7	21.55	3
Dtr	13.85	8	-----	---
P_06	13.84	9	16.47	10
P_G	13.38	10	12.65	15
T_A	12.63	11	16.63	9
P_07	12.25	12	11.01	16
P_05	11.98	13	9.99	20
P_08	11.98	14	14.93	12
P_A	11.97	15	15.21	11
P_W	11.93	16	17.29	8
T_D	11.48	17	10.98	17
P_09	10.18	18	10.53	19
T_09	9.09	19	-----	---
T_G	7.72	20	-----	---
T_01	6.33	21	10.73	18
Sl	-----	---	19.67	6
Solar	-----	---	14.55	13
T_06	-----	---	9.77	21

**Table 3.3: Spatial autocorrelation and its level of significance for Tree Size Class Diversity ( $H_d$ ) and Tree Species Diversity ( $H_s$ ) within the Alaskan boreal forest, and for the residuals of the random forest analysis (RFA) models used for predicting  $H_d$  and  $H_s$  from this dataset.**

	Moran's $I$	$P$ -Value	Geary's $C$	$P$ -Value
Tree Size Class Diversity ( $H_d$ )	0.2642	<0.001	0.7188	<0.001
Tree Species Diversity ( $H_s$ )	0.2704	<0.001	0.7178	<0.001
Residuals of RFA model for $H_d$	-0.2266	0.9999	1.2086	0.9999
Residuals of RFA model for $H_s$	-0.2957	0.9999	1.2868	0.9999

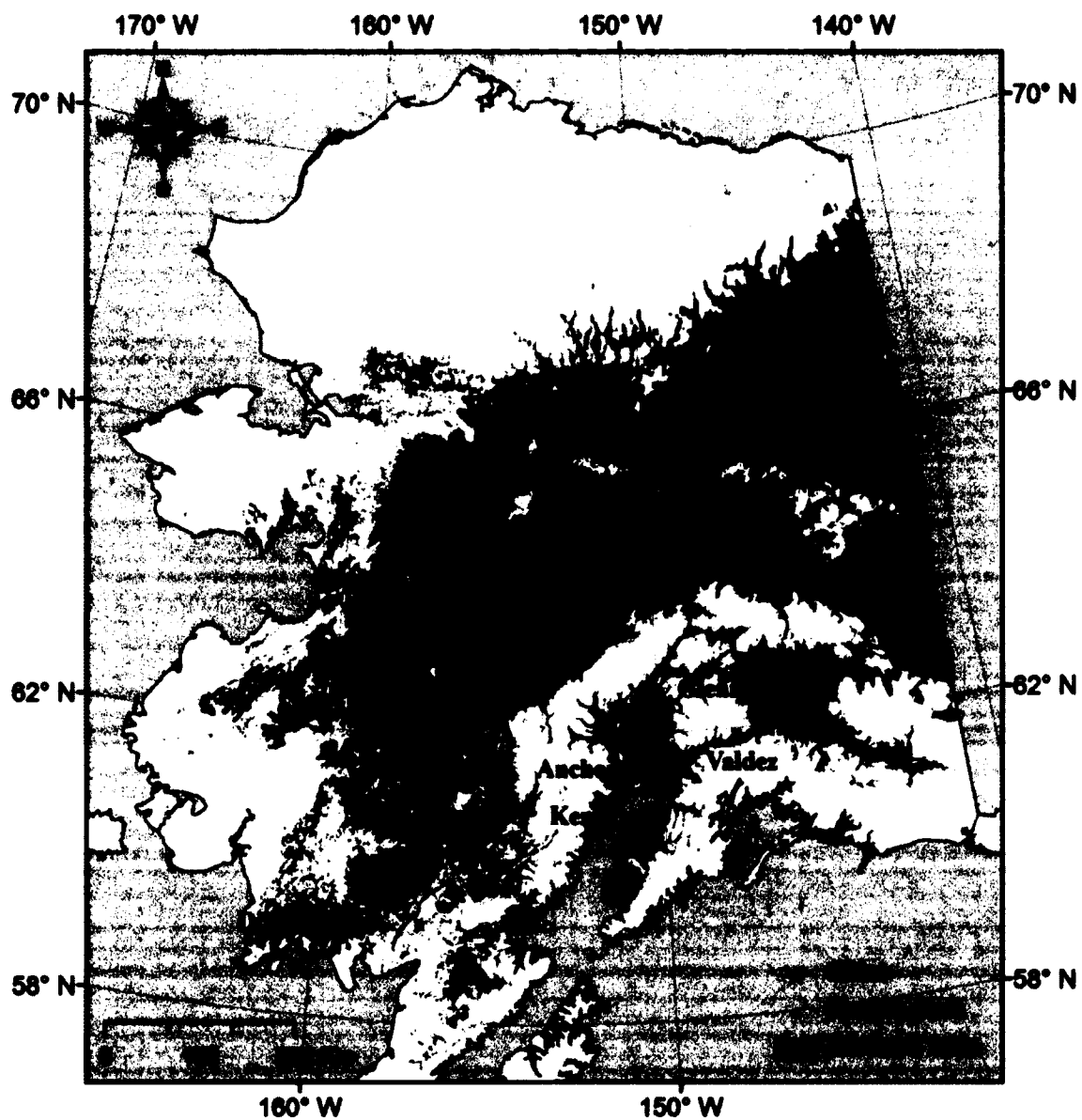


Figure 3.1: Geographic distribution of the 704 Permanent Sample Plots (in triangles) within the Alaskan boreal forest (Ruefenacht et al., 2008).

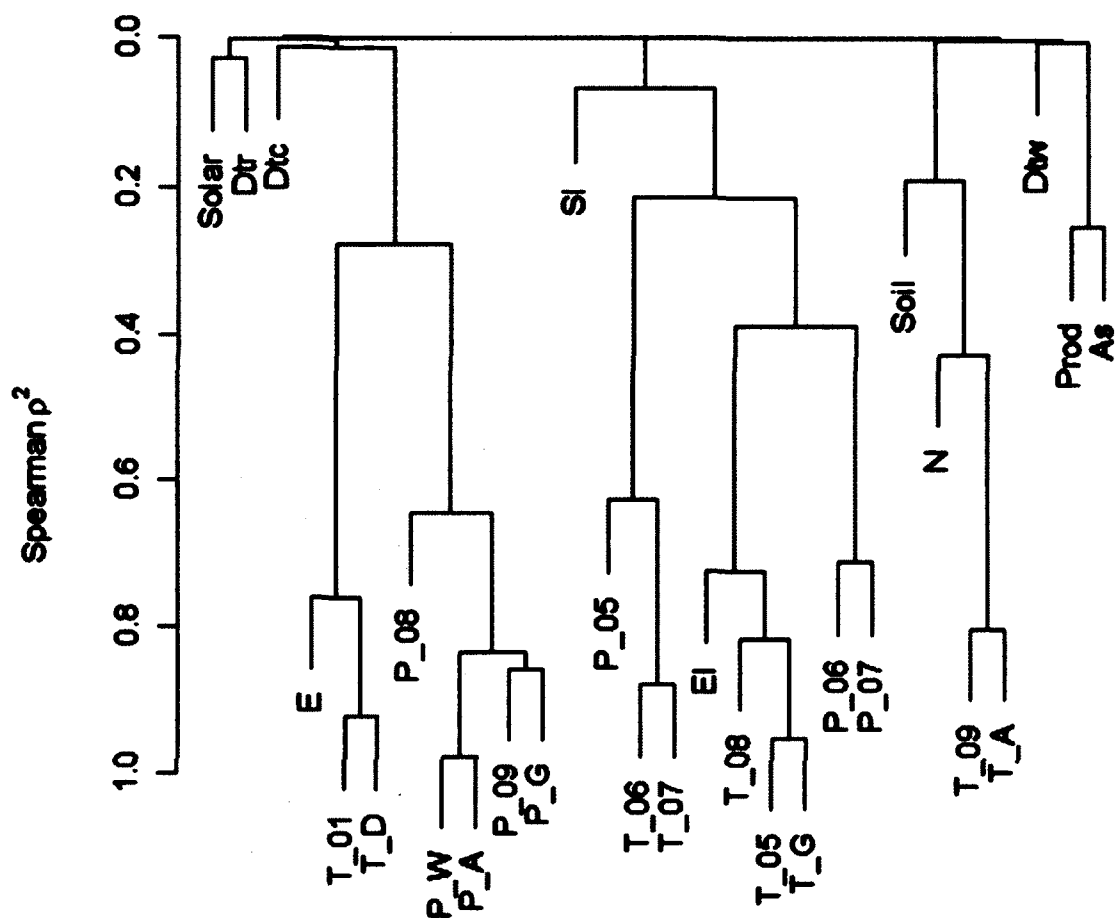


Figure 3.2: Hierarchical clustering using squared Spearman correlation ( $p^2$ ) of the environmental variable similarities.

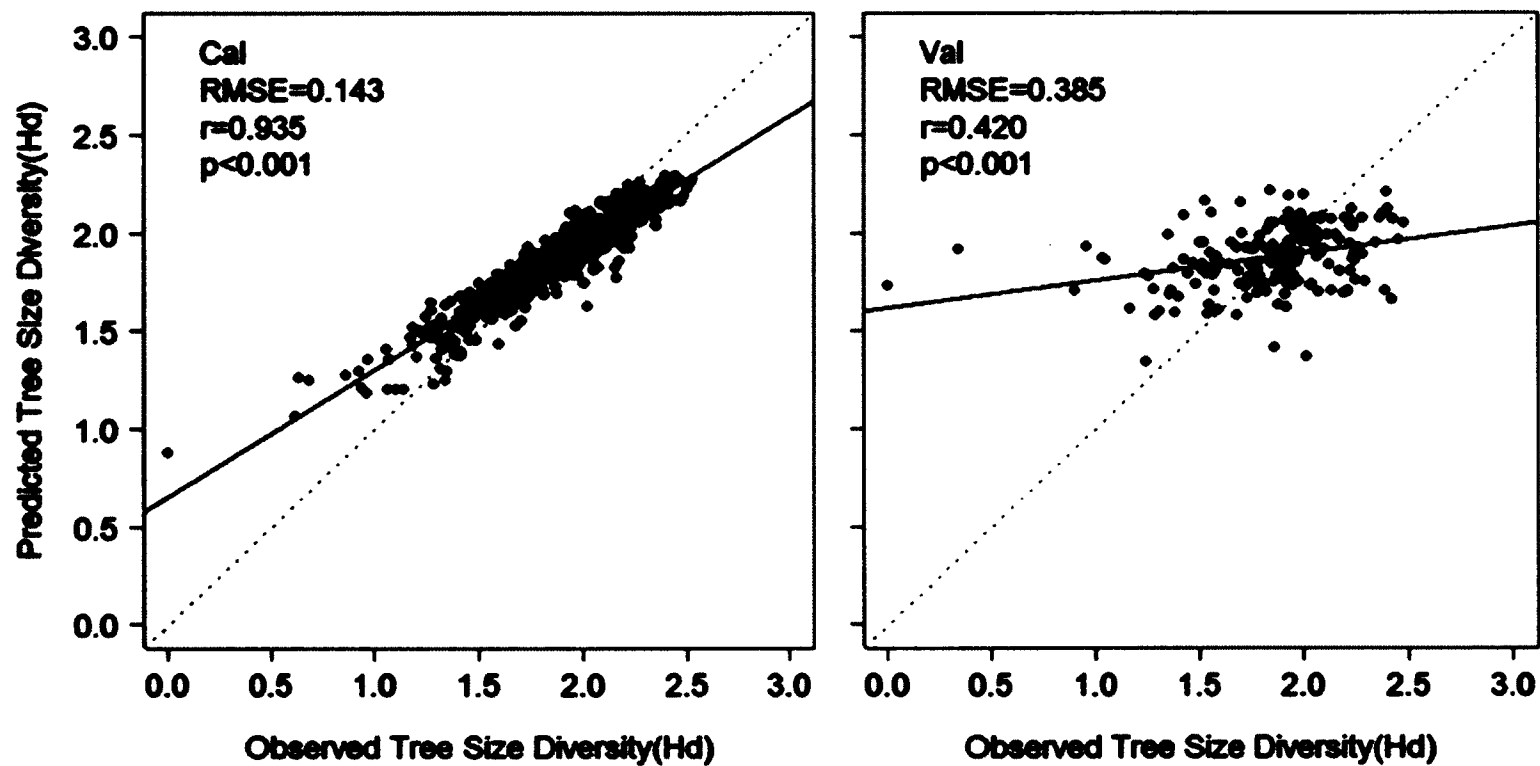


Figure 3.3: Application of the RFA model for tree size-class diversity ( $H_d$ ) to the calibration data set (Cal) and the validation data set (Val).



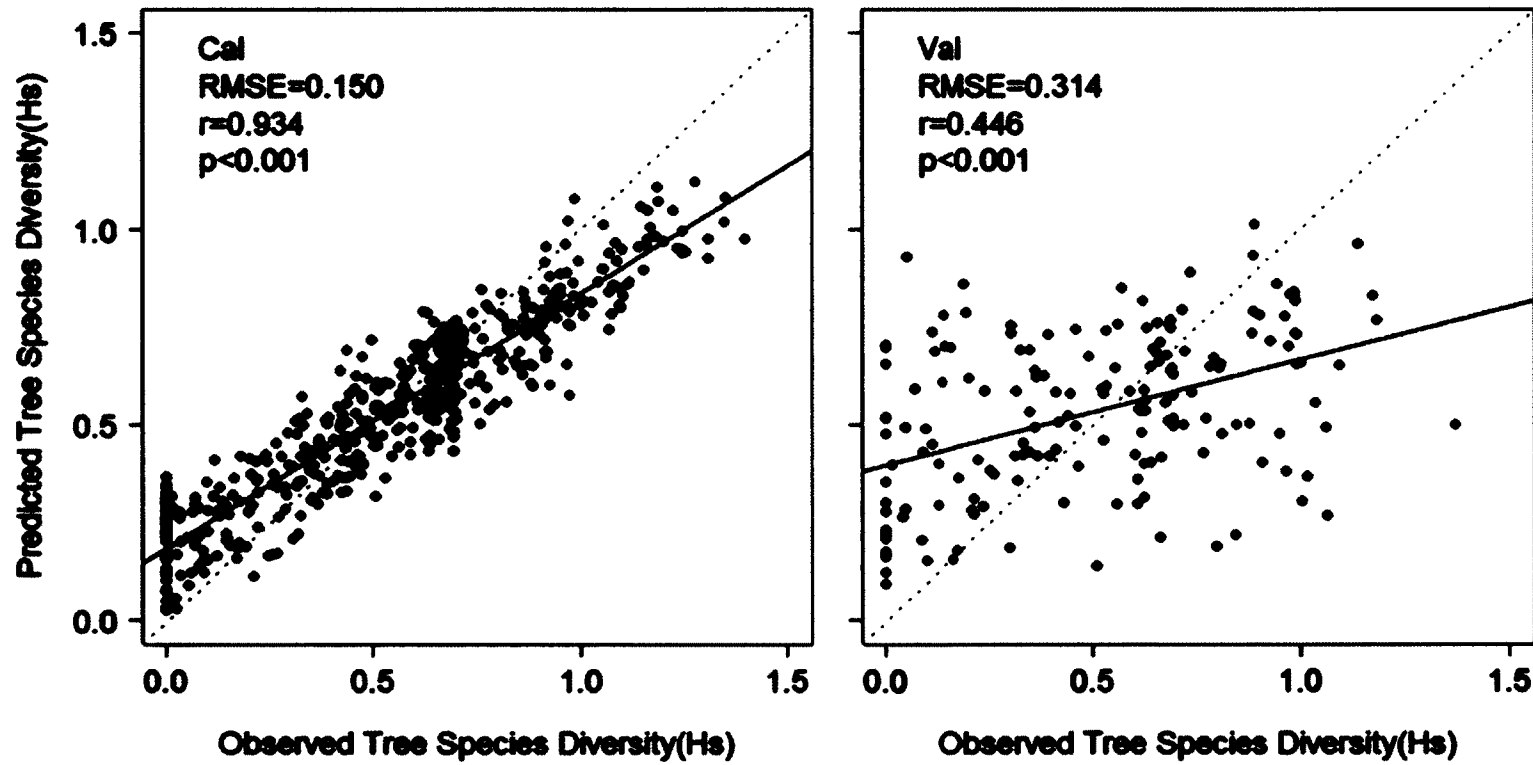


Figure 3.4: Application of the RFA model for tree species diversity ( $H_s$ ) to the calibration data set (Cal) and the validation data set (Val).

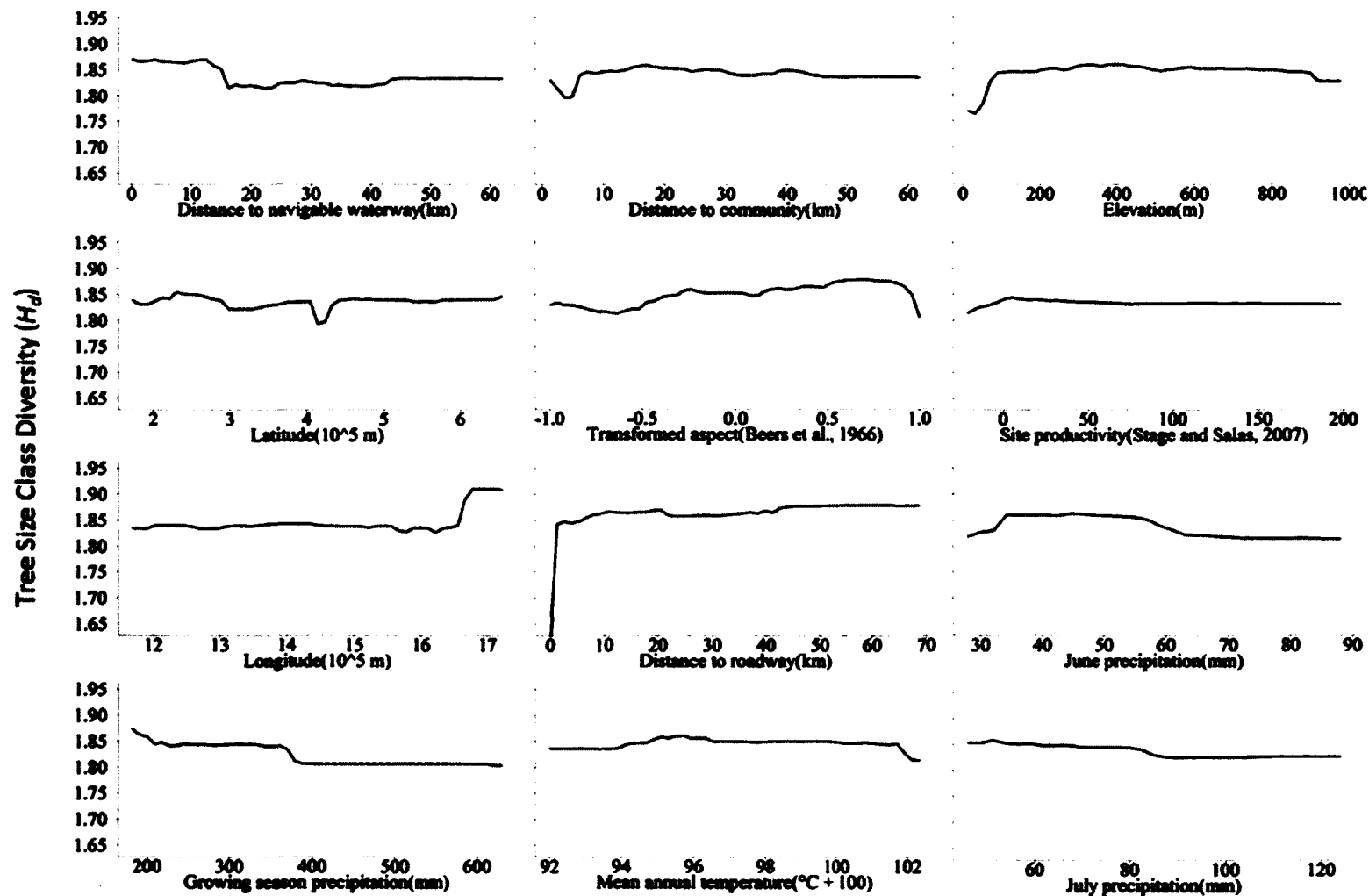


Figure 3.5: Partial plots representing the marginal effects of the twelve most influential variables from the RFA model on the estimates of tree size-class diversity ( $H_d$ ) while averaging out the effect of all the other variables.

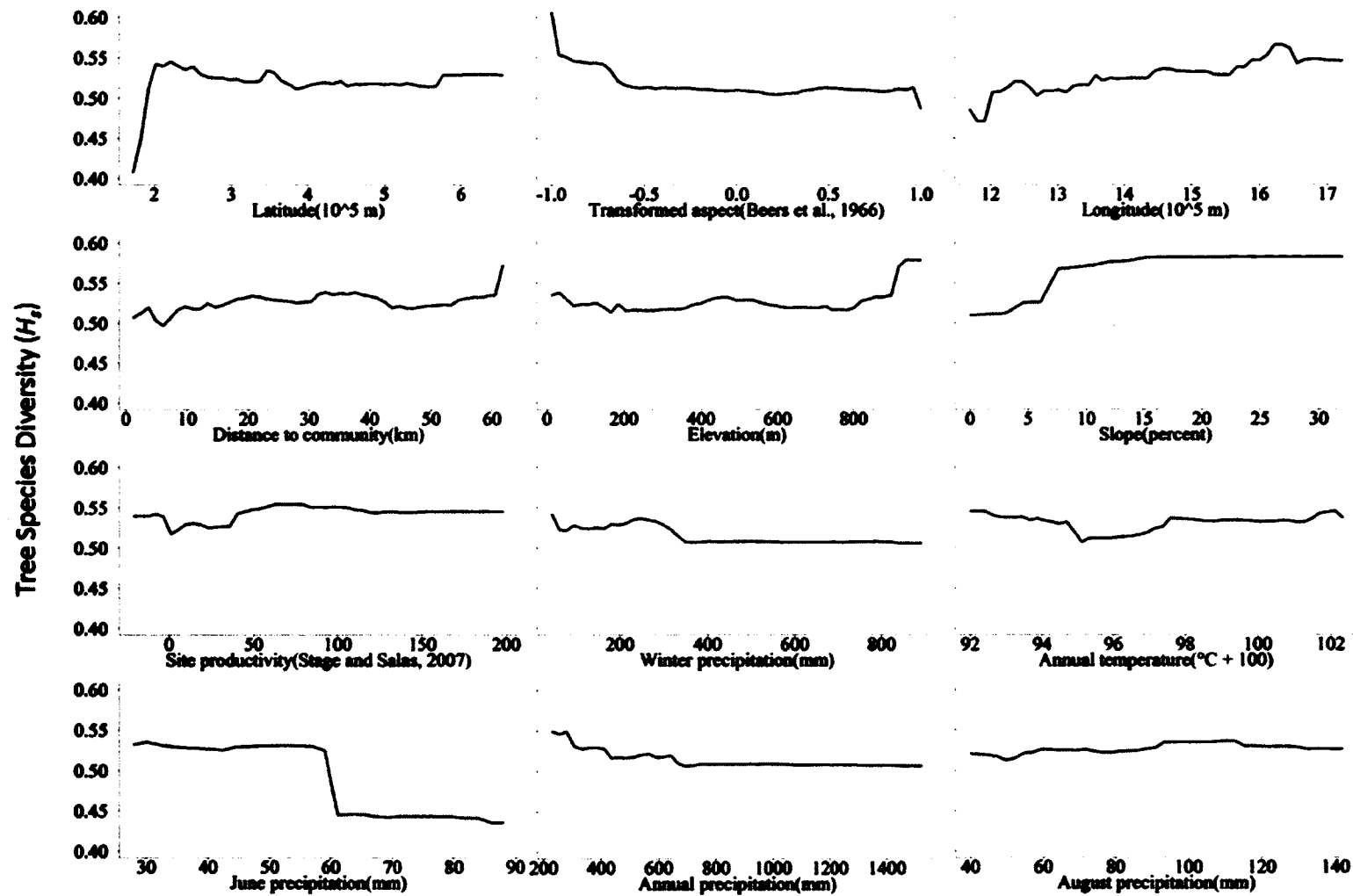


Figure 3.6: Partial plots representing the marginal effects of the twelve most influential variables from the RFA model on the estimates of tree species diversity ( $H_s$ ) while averaging out the effect of all the other variables.

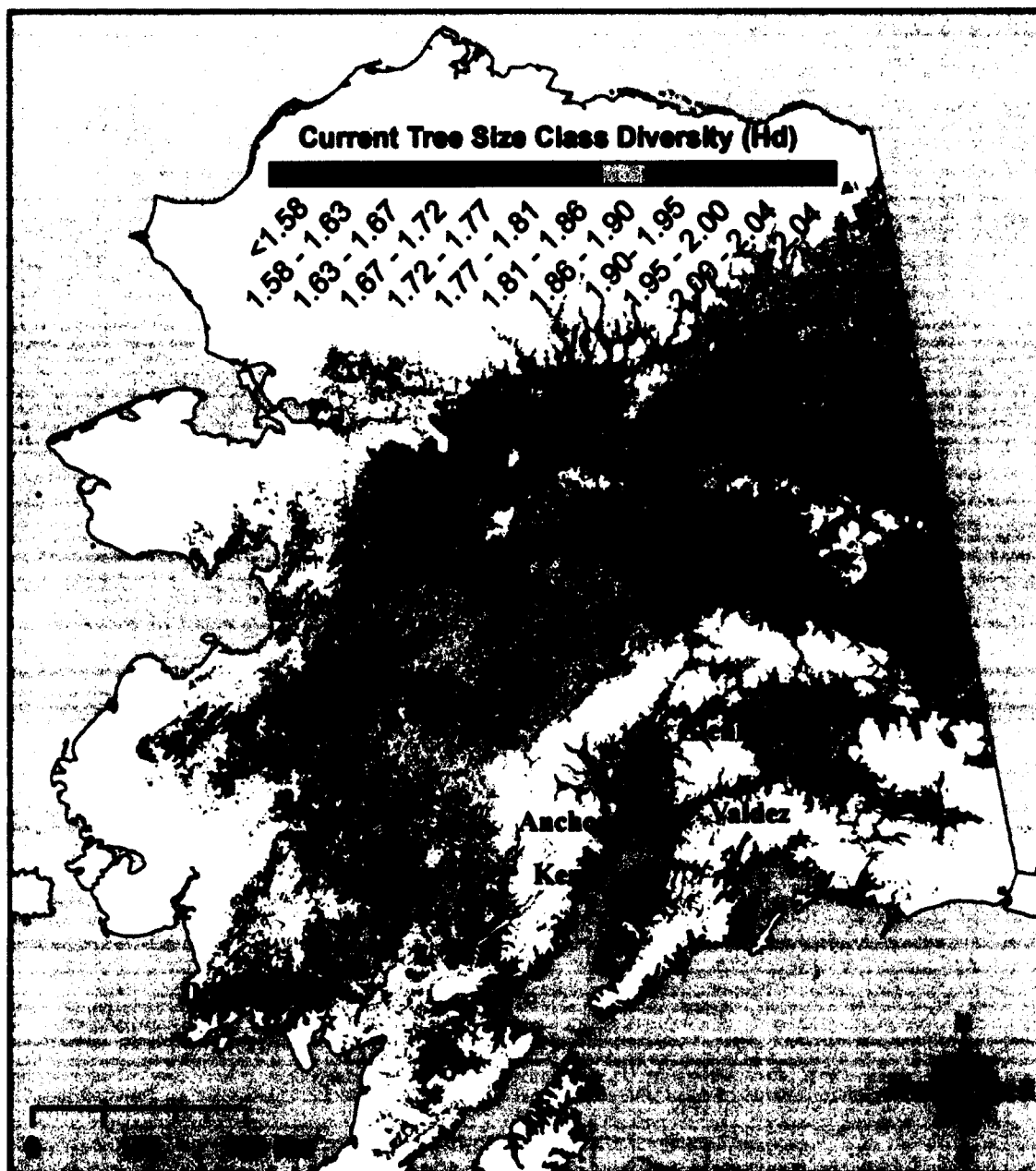


Figure 3.7: Current predicted tree size-class diversity ( $H_d$ ) in the Alaskan boreal forest using the RFA model.

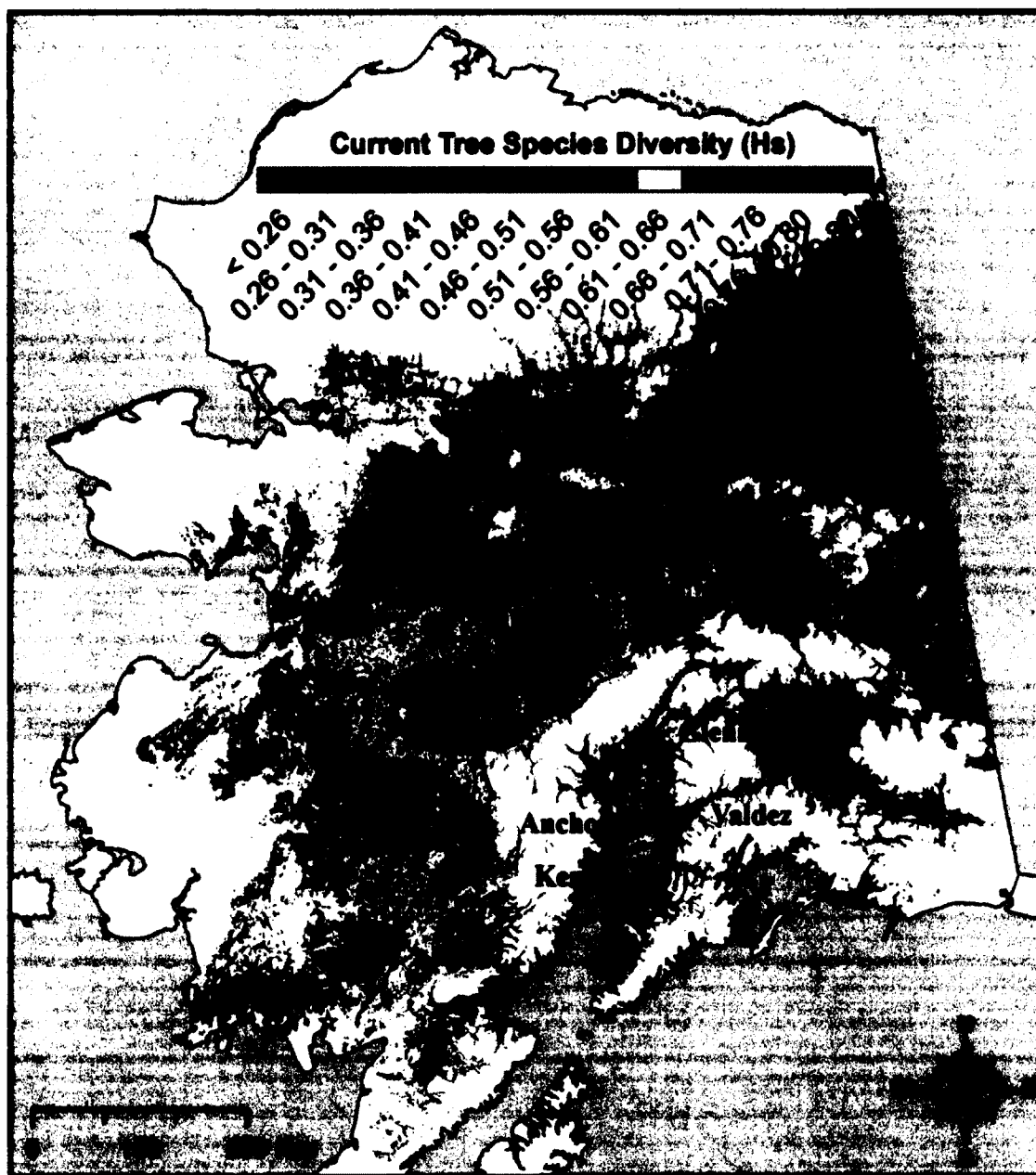


Figure 3.8: Current predicted tree species diversity ( $H_s$ ) in the Alaskan boreal forest using the RFA model.

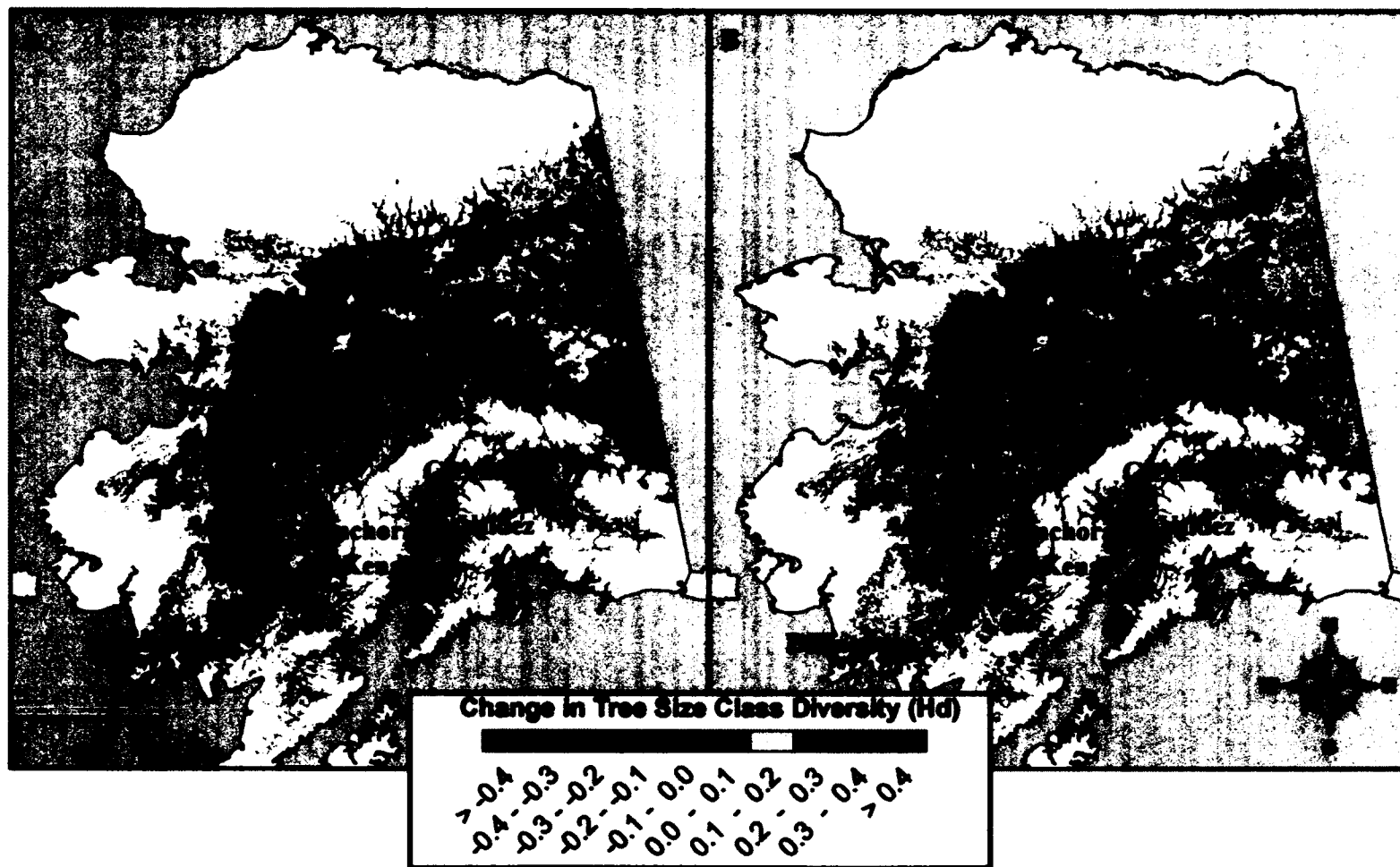


Figure 3.9: Predicted changes in tree size-class diversity ( $H_d$ ) from its current level (Figure 3.6) to the year 2050 for a null climate change scenario (A) and the IPCC A1B scenario (B).

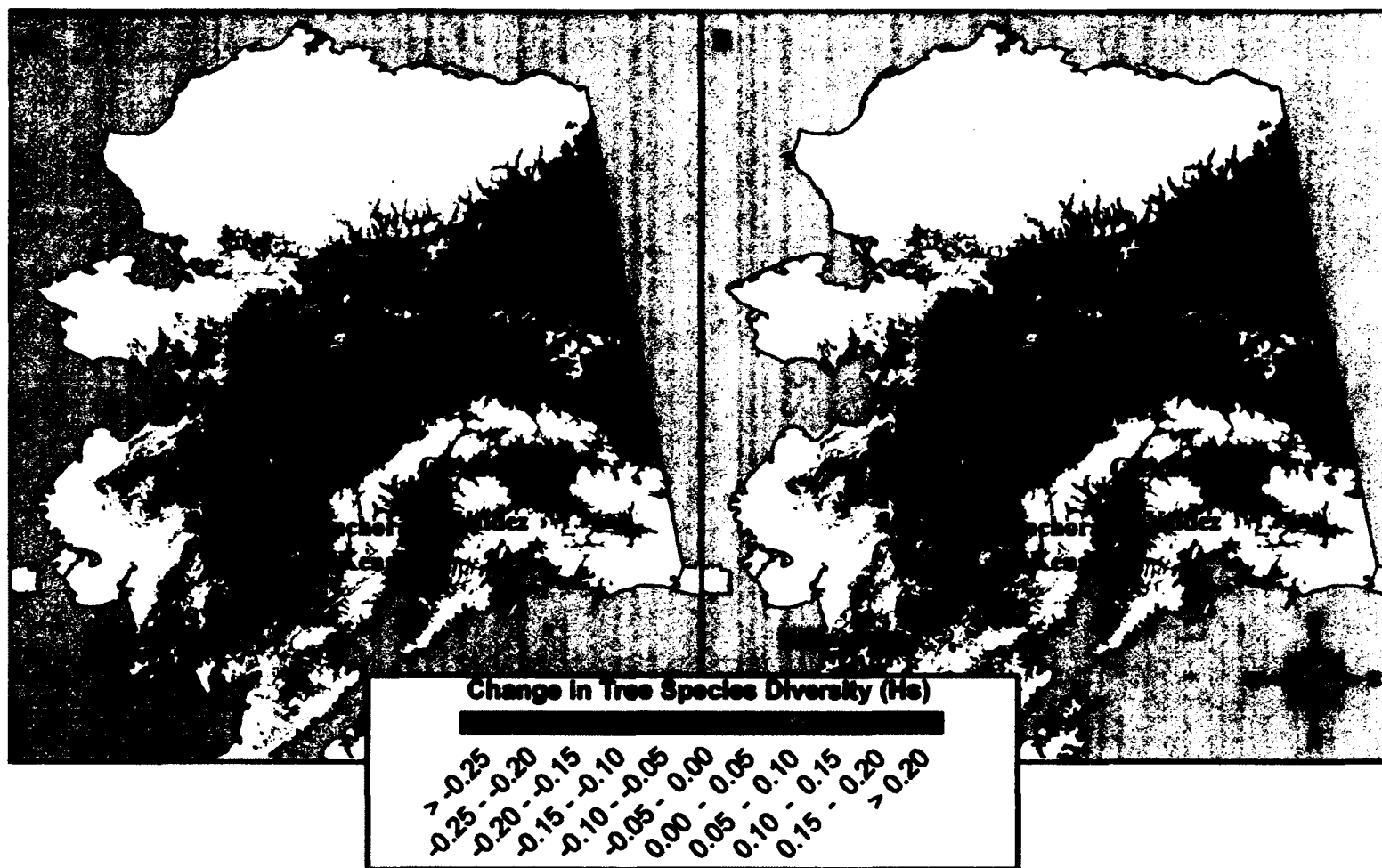


Figure 3.10: Predicted changes in tree species diversity ( $H_s$ ) from its current level (Figure 3.7) to the year 2050 for a null climate change scenario (A) and the IPCC A1B scenario (B).

### Appendix 3.1: R code used in the analysis for Chapter 3

```
#Load library
library(geoR)
library(MASS)
library(spdep)
library(RANN)
library(gstat)
library(sp)
library(maptools)
library(RColorBrewer)
library(hier.part)
library(survival)
library(VGAM)
library(rpart)
library(tree)
library(randomForest)
library(ModelMap)
library(party)
library(Boruta)
library(Hmisc)

#load the data
# set seed
set.seed(71541)
#Training set data (CAFI and WAIN datasets)
mydata=read.csv(file.choose(),header=T)

#extract predictor variables
env =
cbind(mydata$E,mydata$N,mydata$Solar,mydata$T_01,mydata$T_05,
mydata$T_06,mydata$T_07,mydata$T_08,mydata$T_09,mydata$T_G,m
ydata$T_D,mydata$T_A,mydata$P_05,mydata$P_06,mydata$P_07,myda
ta$P_08,mydata$P_09,mydata$P_G,mydata$P_W,mydata$P_A,mydata$P
rod,mydata$SI,mydata$As,mydata$EI,mydata$Dtc,mydata$Dtr,mydata$D
tw,mydata$Soil)
env = as.data.frame(env)

names(env)=c("E","N","Solar","T_01","T_05","T_06","T_07","T_08","T
_09","T_G","T_D","T_A","P_05","P_06","P_07","P_08","P_09","P_G","
P_W","P_A","Prod","SI","As","EI","Dtc","Dtr","Dtw","Soil")
```



```

#hierarchical cluster analysis on variables
env = as.matrix(env)
v=varclus(env)
print(round(v$sim,2))
par(mfrow=c(5.1,4.1,4.1,2.1))
plot(v)
#subset for calibration and validation data
select <- sample(1:nrow(mydata), round(nrow(mydata)*.25,0), replace =
FALSE)
dataval=mydata[select,]
datacal=mydata[-select,]
#subset data for 40 year values only
mydata$Hd40true=mapply(is.na,mydata$Hd40)
mydata40=subset(mydata,Hd40true==FALSE,select=-Hd40true)

#full grid data
datacurrent <- read.csv(file.choose(),header=T)
datacurrent$Soil=as.factor(datacurrent$Soil)

names(datacurrent)[1:26]=c("E","N","EI","SI","As1","Soil","Dtc","P_05",
"P_06","P_07","P_08","P_09","P_G","P_A","T_01","T_05","T_06","T_0
7","T_08","T_09","T_G","T_D","T_A","As","Prod","P_W")

#load functions used
#calculate r2
r2 <- function(obs,pred) {
  SSE <- sum((obs-pred)^2)
  SST <- sum((obs-mean(obs))^2)
  return((1-SSE/SST))
}
#calculate root mean square error (rmse)
rmse <- function(x,y) { sqrt(mean((x-y)^2,na.rm=TRUE)) }
#calculate mean square error (mse)
mse <- function(x,y) { (mean((x-y)^2,na.rm=TRUE)) }
#calculate bias
bias <- function(x,y) { (mean((x-y),na.rm=TRUE)) }
#calculate mean absolute error (MAE)
mae <- function(x,y) { (mean(abs(x-y),na.rm=TRUE)) }
#convert coordinates data to matrix to be used for neighborhood structure
xy=as.matrix(as.data.frame(cbind(mydata$E,mydata$N)))
#identify neighbors
k1a=knearneigh(xy,k=10,longlat=NULL,RANN=TRUE)
all.linked1=max(unlist(nbdists(k1a,xy,longlat=NULL)))
dist=19740

```

```

    nb1=dnearneigh(xy, d1=0, d2=dist, row.names=NULL,longlat=NULL)
    summary(nb1,xy,longlat=NULL)
#plot neighbor data
    plot(xy)
    plot(nb1,xy, add=TRUE,col="red", lty=2, cex=1)
#convert neighbor matrix to weight matrix
    dlist1=nbdists(nb1,xy)
    idlist1=lapply(dlist1,function(x)1/x)
    nbw1=nb2listw(nb1,glist=idlist1,style="W",zero.policy=TRUE)
    summary(nbw1)
#tests for autocorrelation in RF final model for Hd
    lmnetmoran=moran.mc(resid_Hd_final,nbw1,zero.policy=TRUE,nsim=9999)
    lmnetmoran
    lmnetgeary=geary.mc(resid_Hd_final,nbw1,zero.policy=TRUE,nsim=9999)
    lmnetgeary
#tests for autocorrelation in RF final model for Hs
    lmnetmoran=moran.mc(resid_Hs_Final,nbw1,zero.policy=TRUE,nsim=9999)
    lmnetmoran
    lmnetgeary=geary.mc(resid_Hs_Final,nbw1,zero.policy=TRUE,nsim=9999)
    lmnetgeary
#tests for autocorrelation for Hd
    moranHd=moran.mc(mydata$Hd,nbw1,zero.policy=TRUE,nsim=9999)
    moranHd
    gearyHd=geary.mc(mydata$Hd,nbw1,zero.policy=TRUE,nsim=9999)
    gearyHd
    print(sp.correlogram(nb1, mydata$Hd,order=15, method="I",
    zero.policy=TRUE))
#tests for autocorrelation for Hs
    moranHs=moran.mc(mydata$Hs,nbw1,zero.policy=TRUE,nsim=9999)
    moranHs
    gearyHs=geary.mc(mydata$Hs,nbw1,zero.policy=TRUE,nsim=9999)
    gearyHs
    print(sp.correlogram(nb1, mydata$Hs,order=15, method="I",
    zero.policy=TRUE))
#tests for autocorrelation for T_A
    moranT_A=moran.mc(mydata$T_A,nbw1,zero.policy=TRUE,nsim=9999)
    moranT_A
    gearyT_A=geary.mc(mydata$T_A,nbw1,zero.policy=TRUE,nsim=9999)
    gearyT_A
    print(sp.correlogram(nb1, mydata$T_A,order=15, method="I",
    zero.policy=TRUE))
#tests for autocorrelation for P_A
    moranP_A=moran.mc(mydata$P_A,nbw1,zero.policy=TRUE,nsim=9999)
    moranP_A

```

```

gearyP_A=geary.mc(mydata$P_A,nbw1,zero.policy=TRUE,nsim=9999)
gearyP_A
plot(sp.correlogram(nb1, mydata$P_A,order=15, method="corr",
zero.policy=TRUE))
print(sp.correlogram(nb1, mydata$P_A,order=15, method="I",
zero.policy=TRUE))

```

#### #binary tree regression

```

names(mydata)
xnam <- names(mydata)[4:27]
xnam_utm <- names(mydata)[1:27]
(fm1a_Hd <- as.formula(paste("Hd ~ ", paste(xnam, collapse= "+"))))
(fm1a_Hs <- as.formula(paste("Hs ~ ", paste(xnam, collapse= "+"))))
(fm1a_Hd_utm <- as.formula(paste("Hd ~ ", paste(xnam_utm, collapse= "+"))))
(fm1a_Hs_utm <- as.formula(paste("Hs ~ ", paste(xnam_utm, collapse= "+"))))
model_Hd<-tree(fm1a_Hd_utm,data=mydata)
print(model_Hd)
plot(model_Hd)
text(model_Hd)
model_Hs<-tree(fm1a_Hs,data=mydata)
print(model_Hs)
plot(model_Hs)
text(model_Hs)

```

#### #stepwise model selection

```

null_Hd=lm(Hd~1,data=mydata)
full_Hd=lm(fm1a_Hd_utm,data=mydata)
step_Hd=step(full_Hd,data=mydata,direction="both",k=log(nrow(mydata)))
null_Hs=lm(Hs~1,data=mydata)
full_Hs=lm(fm1a_Hs,data=mydata)
step_Hs=step(full_Hs,data=mydata,direction="both",k=log(nrow(mydata)))

```

#### # lm models using stepwise predictors

```

model_Hd=lm(Hd ~ N + E + T_01 + T_09 + T_G + T_D +
             T_A + P_05 + P_06 + P_07 + P_08 +
             P_09 + P_G + P_A + El + As + Dtc +
             Soil,data = mydata)
summary(model_Hd)
model_Hs=lm(Hs ~ El + Sl + T_07 + P_08 + P_A, data = mydata)
summary(model_Hs)

```

# GAM model using stepwise variables

```
library(mgcv)
Tmodel_Hd=gam(Hd ~ te(E,N) + T_07 + P_06 + P_07, data = mydata)
summary(Tmodel_Hd)
Tmodel_Hs=gam(Hs ~ te(E,N) + El + Sl + P_07 + P_08 + P_A, data = mydata)
summary(Tmodel_Hs)
```

#Using Random forest to determine best variables for Hd with variables which are not subject to change + climate change variables...

```
valfinal_Hd <- Boruta(Hd ~ N + E + T_01 + T_05 + T_06 + T_07 +
  T_08 + T_09 + T_G + T_D + T_A + P_05 +
  P_06 + P_07 + P_08 + P_09 + P_G + P_A + P_W +
  Solar + El + As + Sl + Dtc + Dtw + Dtr + Soil +
  Prod,data=mydata,maxRuns = 500,doTrace=2)

print(valfinal_Hd)
plot(valfinal_Hd,whichRand=c(FALSE,FALSE,FALSE),cex.lab=1.8,
cex.axis=1.8)
getConfirmedFormula(valfinal_Hd)
stats<-attStats(valfinal_Hd)
print(stats)
```

#analyze final model using calibration and validation data sets

```
final_Rand_Hd <- randomForest(Hd ~ N + E + T_01 + T_09 + T_G +
  T_D + T_A + P_05 + P_06 + P_07 + P_08 + P_09 + P_G + P_A + P_W +
  El + As + Dtc + Dtw + Dtr + Prod,data=mydata,ntree=500, mtry=6,
  importance=TRUE)

round(importance(final_Rand_Hd,type=1),2)
predict_final_Hd=predict(final_Rand_Hd,mydata)
rmse(predict_final_Hd,mydata$Hd)
bias(predict_final_Hd,mydata$Hd)
cor.test(predict_final_Hd,mydata$Hd,method="pearson")
mae(predict_final_Hd,mydata$Hd)
mean(final_Rand_Hd$mse)
mean(final_Rand_Hd$rsq)
resid_Hd_final=predict_final_Hd - mydata$Hd
plot(resid_Hd_final~predict_final_Hd)
```

#Calibration set analysis

```
final_Rand_Hd_cal <- randomForest(Hd ~ N + E + T_01 + T_09 + T_G +
  T_D + T_A + P_05 + P_06 + P_07 + P_08 + P_09 + P_G + P_A + P_W +
  El + As + Dtc + Dtw + Dtr + Prod,data=atacal,ntree=500,mtry=6,
  importance=TRUE)
```

```

round(importance(final_Rand_Hd_cal,type=1),2)
predict_final_Hd_cal=predict(final_Rand_Hd_cal,datacal)
rmse(predict_final_Hd_cal,datacal$Hd)

cor.test(predict_final_Hd_cal,datacal$Hd,method="pearson")

```

# Plot results from RF model

```

par(mar = c(5.1,6.1,4.1,2.1)+0.1, mgp=c(3.5,1,0))

plot(predict_final_Hd_cal~datacal$Hd,xlab="Observed Tree Size
Diversity(Hd)",ylab="",xlim=c(0,3),ylim=c(0,3),cex.axis=1.5,cex.lab=1.5,
las=1,pch=19)
mtext("Predicted Tree Size Diversity(Hd)",cex=1.5,side=2,line=4)

```

```

myline.fit.cal <- lm(predict_final_Hd_cal~datacal$Hd)
summary(myline.fit.cal)
abline(myline.fit.cal,col="RED",cex=1.5,lwd=2)
abline(a=0,b=1,lty="dotted",col="gray70",cex=1.5,lwd=2)
legend("topleft",
c("Cal", "RMSE=0.146", "r=0.942", "p<0.001"),col=c("black"), bty =
"n",cex=1.5)

```

```

predict_final_Hd_val=predict(final_Rand_Hd_cal,dataval)
rmse(predict_final_Hd_cal,dataval$Hd)
cor.test(predict_final_Hd_val,dataval$Hd,method="pearson")

plot(predict_final_Hd_val~dataval$Hd,xlab="Observed Tree Size
Diversity(Hd)",ylab="",xlim=c(0,3),ylim=c(0,3),cex.axis=1.5,cex.lab=1.5,
las=1,pch=19)
mtext("Predicted Tree Size Diversity(Hd)",cex=1.5,side=2,line=4)

```

```

myline.fit.val <- lm(predict_final_Hd_val~dataval$Hd)
summary(myline.fit.val)
abline(myline.fit.val,col="RED",cex=1.5,lwd=2)
abline(a=0,b=1,lty="dotted",col="gray70",cex=1.5,lwd=2)
legend("topleft",
c("Val", "RMSE=0.385", "r=0.445", "p<0.001"),col=c("black"), bty =
"n",cex=1.5)

```

#creation of full model for Hd\_2050 Null Climate

```

names(mydata40)
xnam_utm <- names(mydata40)[c(4:6,10:27)]
(fmla_Hd40 <- as.formula(paste("Hd40 ~ ", paste(xnam_utm, collapse= "+"))))

```

```
#predict for Hd_2050 Null Climate
```

```
Rand_Hd40 <- randomForest(Hd40 ~ Tmean_01 + Tmean_05 + Tmean_06 +
  Tmean_08 + Tmean_09 + T_Diff + Tmean_Ann + Precip_05 + Precip_06 +
  Precip_07 + Precip_08 + Precip_09 + Precip_G + Precip_Ann + El + As + Sl +
  Soil,data=mydata40,importance=TRUE)
```

```
round(importance(Rand_Hd40,type=1),2)
predict_Rand_Hd40=predict(Rand_Hd40,mydata40)
rmse(predict_Rand_Hd40,mydata40$Hd40)
bias(predict_Rand_Hd40,mydata40$Hd40)
cor.test(predict_Rand_Hd40,mydata40$Hd40,method="pearson")
```

```
#RF model diagnostics
```

```
plot(predict_final_Hd_val,dataval$Hd,xlab="Predicted Hd",ylab="Observed Hd",
  xlim=c(0,3),ylim=c(0,3),cex.axis=1.5,cex.lab=1.5,las=1)
```

```
myline.fit <- lm(dataval$Hd~predict_final_Hd_val)
```

```
summary(myline.fit)
abline(myline.fit,col="black",cex=1.5,lwd=2)
abline(a=0,b=1,lty="dotted",col="gray70",cex=1.5,lwd=2)
```

```
plot(resid_Hd,predict_final_Hd_cal)
par(mar = c(9,10,2.5,2.5))
par(family="serif",font=2)
```

```
partialPlot(final_Rand_Hd,mydata,Dtw,main="",ylim=c(1.64,1.95),xlab="Distanc
e to navigable waterway(km)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp =
c(6,2.5,0), bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,Dtc,main="",ylim=c(1.64,1.95),xlab="Distanc
e to community(km)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,El,main="",ylim=c(1.64,1.95),xlab="Elevation
(m)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0), bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,N,main="",ylim=c(1.64,1.95),xlab="Latitude(
10^5 m)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,As,main="",ylim=c(1.64,1.95),xlab="Transfor
med aspect(Beers et al., 1966)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp =
c(6,2.5,0), bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,Prod,main="",ylim=c(1.64,1.95),xlab="Site
productivity(Stage and Salas, 2007)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp =
c(6,2.5,0), bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,E,main="",ylim=c(1.64,1.95),xlab="Longitude
(10^5 m)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,Dtr,main="",ylim=c(1.64,1.95),xlab="Distanc
e to roadway(km)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,P_06,main="",ylim=c(1.64,1.95),xlab="June
precipitation(mm)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,P_G,main="",ylim=c(1.64,1.95),xlab="Growi
ng season precipitation(mm)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp =
c(6,2.5,0), bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,T_A,main="",ylim=c(1.64,1.95),xlab="Mean
annual temperature(°C + 100)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp =
c(6,2.5,0), bty="l",rug=FALSE)
```

```
partialPlot(final_Rand_Hd,mydata,P_07,main="",ylim=c(1.64,1.95),xlab="July
precipitation(mm)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
```

```
#predict values for all locations for current Hd!
```

```
AKpredict_Hd <- predict(final_Rand_Hd,datacurrent)
Hd_spat=cbind(datacurrent$E*100000,datacurrent$N*100000,AKpredict_Hd)
Hd_spat=as.data.frame(Hd_spat)
names(Hd_spat)=c("E", "N", "Hd")
write.csv(Hd_spat,"Hdcurrentvalues.csv",row.names=FALSE)
```

```
#predict values for all locations for Hd40!
```

```
AKpredict_Hd40 <- predict(Rand_Hd40,datacurrent)
Hd40_spat=cbind(datacurrent$E,datacurrent$N,AKpredict_Hd40)
Hd40_spat=as.data.frame(Hd40_spat)
names(Hd40_spat)=c("E", "N", "Hd40")
write.csv(Hd40_spat,"Hd40values.csv",row.names=FALSE)
```

```

#predict values for all locations for 2050 Hd40!
AKpredict_Hd40_2050 <- predict(Rand_Hd40,data2050)
Hd40_spat_2050=cbind(data2050$E,data2050$N,AKpredict_Hd40_2050)
Hd40_spat_2050=as.data.frame(Hd40_spat_2050)
names(Hd40_spat_2050)=c("E","N","Hd40")
write.csv(Hd40_spat_2050,"Hd40values2050.csv",row.names=FALSE)

#Using Random forest to determine best variables for Hs for climate change variables
valfinal_Hs <- Boruta(Hs ~ N + E + T_01 + T_05 + T_06 + T_07 + T_08 +
                      T_09 + T_G + T_D + T_A + P_05 + P_06 +
                      P_07 + P_08 + P_09 + P_G + P_A + P_W +
                      Solar + El + As + Sl + Dtc + Dtw + Dtr + Soil +
                      Prod, data=mydata,maxRuns = 500,doTrace=2)

print(valfinal_Hs)
plot(valfinal_Hs,whichRand=c(FALSE,FALSE,FALSE),cex.lab=1.8,
cex.axis=1.8)
getConfirmedFormula(valfinal_Hs)
stats<-attStats(valfinal_Hs)
print(stats)

#calibrate and validate model Boruta model
Final_Rand_Hs <- randomForest(Hs ~ N + E + T_01 + T_06 +
                              T_D + T_A + P_05 + P_06 + P_07 +
                              P_08 + P_09 + P_G + P_A + P_W +
                              Solar + El + As + Sl + Dtc + Dtw + Prod,
                              data=mydata,ntree=500,mtry=7,importance=TRUE)

round(importance(Final_Rand_Hs,type=1),2)
predict_Final_Hs=predict(Final_Rand_Hs,mydata)
rmse(predict_Final_Hs,mydata$Hs)
cor.test(predict_Final_Hs,mydata$Hs,method="pearson")
bias(predict_Final_Hs,mydata$Hs)
mae(predict_Final_Hs,mydata$Hs)
resid_Hs_Final=predict_Final_Hs - mydata$Hs
plot(resid_Hs_Final~predict_Final_Hs)

Final_Rand_Hs_cal <- randomForest(Hs ~ N + E + T_01 + T_06 +
                                  T_D + T_A + P_05 + P_06 + P_07 +
                                  P_08 + P_09 + P_G + P_A + P_W +
                                  Solar + El + As + Sl + Dtc + Dtw + Prod,
                                  data=atacal,ntree=500,mtry=7,importance=TRUE)

round(importance(Final_Rand_Hs_cal,type=1),2)
predict_Final_Hs_cal=predict(Final_Rand_Hs_cal,atacal)
rmse(predict_Final_Hs_cal,atacal$Hs)
cor.test(predict_Final_Hs_cal,atacal$Hs,method="pearson")

```



```

bias(predict_Final_Hs_cal,datacal$Hs)
resid_Hs_Final_cal=predict_Final_Hs_cal - datacal$Hs
plot(resid_Hs_Final_cal~predict_Final_Hs_cal)

```

```

par(mar = c(5.1,6.1,4.1,2.1)+0.1, mgp=c(3.5,1,0))

```

```

plot(predict_Final_Hs_cal~datacal$Hs,xlab="Observed Tree Species
Diversity(Hs)",ylab="",xlim=c(0,1.5),ylim=c(0,1.5),cex.axis=1.5,cex.lab=1.5,las=
1,pch=19)mtext("Predicted Tree Species Diversity(Hs)",cex=1.5,side=2,line=4)

```

```

myline.fit.cal <- lm(predict_Final_Hs_cal~datacal$Hs)
summary(myline.fit.cal)
abline(myline.fit.cal,col="RED",cex=1.5,lwd=2)
abline(a=0,b=1,lty="dotted",col="gray70",cex=1.5,lwd=2)
legend("topleft", c("Cal", "RMSE=0.143", "r=0.946", "p<0.001"),col=c("black"),
bty = "n",cex=1.5)

```

```

predict_Final_Hs_val=predict(Final_Rand_Hs_cal,dataval)
rmse(predict_Final_Hs_val,dataval$Hs)
cor.test(predict_Final_Hs_val,dataval$Hs,method="pearson")
bias(predict_Final_Hs_val,dataval$Hs)
resid_Hs_Final_val=predict_Final_Hs_val - dataval$Hs
plot(resid_Hs_Final_val~predict_Final_Hs_val)

```

```

plot(predict_Final_Hs_val~dataval$Hs,xlab="Observed Tree Species
Diversity(Hs)",ylab="",xlim=c(0,1.5),ylim=c(0,1.5),cex.axis=1.5,cex.lab=1.5,las=
1,pch=19)mtext("Predicted Tree Species Diversity(Hs)",cex=1.5,side=2,line=4)

```

```

myline.fit.val <- lm(predict_Final_Hs_val~dataval$Hs)
summary(myline.fit.val)
abline(myline.fit.val,col="RED",cex=1.5,lwd=2)
abline(a=0,b=1,lty="dotted",col="gray70",cex=1.5,lwd=2)
legend("topleft", c("Val", "RMSE=0.314", "r=0.446", "p<0.001"),col=c("black"),
bty = "n",cex=1.5)

```

#predict values for all locations for Hs40!

```

AKpredict_Hs40 <- predict(Final_Rand_Hs,datacurrent)
Hs40_spat=cbind(datacurrent$E,datacurrent$N,AKpredict_Hs40)
Hs40_spat=as.data.frame(Hs40_spat)
names(Hs40_spat)=c("E","N","Hs40")
write.csv(Hs40_spat,"Hs40values.csv",row.names=FALSE)

```

```

#predict values for all locations for 2050 Hs40!
  AKpredict_Hs40_2050 <- predict(Rand_Hs40,data2050)
  Hs40_spat_2050=cbind(data2050$E,data2050$N,AKpredict_Hs40_2050)
  Hs40_spat_2050=as.data.frame(Hs40_spat_2050)
  names(Hs40_spat_2050)=c("E","N","Hs40")
  write.csv(Hs40_spat_2050,"Hs40values2050.csv",row.names=FALSE)

# partial probability plots for the final model for Hs
  partialPlot(Final_Rand_Hs,mydata,N,main="",ylim=c(.40,.60),xlab="Latitude(10
^5 m)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0), bty="l",rug=FALSE)

  partialPlot(Final_Rand_Hs,mydata,As,main="",ylim=c(.40,.60),xlab="Transform
ed aspect(Beers et al., 1966)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
  partialPlot(Final_Rand_Hs,mydata,E,main="",ylim=c(.40,.60),xlab="Longitude(1
0^5 m)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)

  partialPlot(Final_Rand_Hs,mydata,Dtc,main="",ylim=c(.40,.60),xlab="Distance
to community(km)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)

  partialPlot(Final_Rand_Hs,mydata,El,main="",ylim=c(.40,.60),xlab="
Elevation(m)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)

  partialPlot(Final_Rand_Hs,mydata,Sl,main="",ylim=c(.40,.60),xlab="
Slope(percent)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)

  partialPlot(Final_Rand_Hs,mydata,Prod,main="",ylim=c(.40,.60),xlab="Site
productivity(Stage and Salas, 2007)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp =
c(6,2.5,0), bty="l",rug=FALSE)

  partialPlot(Final_Rand_Hs,mydata,P_W,main="",ylim=c(.40,.60),xlab="Winter
precipitation(mm)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)

  partialPlot(Final_Rand_Hs,mydata,T_A,main="",ylim=c(.40,.60),xlab="Annual
temperature(°C + 100)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)

```

```
partialPlot(Final_Rand_Hs,mydata,P_06,main="",ylim=c(.40,.60),xlab="June
precipitation(mm)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
```

```
partialPlot(Final_Rand_Hs,mydata,P_A,main="",ylim=c(.40,.60),xlab="Annual
precipitation(mm)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
```

```
partialPlot(Final_Rand_Hs,mydata,P_08,main="",ylim=c(.40,.60),xlab="August
precipitation(mm)",cex.lab=5,cex.axis=5,las=1,lwd=5,mgp = c(6,2.5,0),
bty="l",rug=FALSE)
```

```
#predict values for all locations for current Hs!
```

```
AKpredict_Hs <- predict(Final_Rand_Hs,datacurrent)
Hs_spat=cbind(datacurrent$E*100000,datacurrent$N*100000,AKpredict_Hs)
Hs_spat=as.data.frame(Hs_spat)
names(Hs_spat)=c("E","N","Hs")
write.csv(Hs_spat,"Hsvaluescurrent.csv",row.names=FALSE)
```

```
#predict values for all locations for 2050 Hs!
```

```
AKpredict_Hs_2050 <- predict(Final_Rand_Hs,data2050)
Hs_spat_2050=cbind(data2050$E,data2050$N,AKpredict_Hs_2050)
Hs_spat_2050=as.data.frame(Hs_spat_2050)
names(Hs_spat_2050)=c("E","N","Hs")
write.csv(Hs_spat_2050,"Hsvalues2050.csv",row.names=FALSE)
```

## Appendix 3.2: Forest types of the boreal forest of Alaska Raster Dataset Metadata

### Standard Version FGDC-STD-001-1998

#### Identification

##### CITATION

##### CITATION INFORMATION

PUBLICATION DATE 2011-04-06

PUBLICATION TIME 000000

##### TITLE

Forest types of the boreal forest of Alaska

GEOSPATIAL DATA PRESENTATION FORM raster digital data

##### SERIES INFORMATION

SERIES NAME Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions

##### DESCRIPTION

##### ABSTRACT

This raster layer was used in a study that modeled and mapped forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions. The input forest types were created by Ruefenacht et al. 2008 and constituted forest types across the conterminous US and Alaska. They were downloaded from the website of Oak Ridge National Laboratory Distributed Active Archive Center and clipped to the region of interest.

The region of interest for this study was the Alaskan boreal forest which extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains.

The abstract for the paper in which this data was originally published is below for reference:

Title: Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions

Brian Young<sup>1</sup>,

John Yarie<sup>1</sup>,

David Verbyla<sup>1</sup>,

Falk Huettmann<sup>2</sup>,

**F. Stuart Chapin III**

**1Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA**

**2Institute of Arctic Biology, Biology & Wildlife Department – EWHALE lab – University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA**

**3Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA**

**Abstract:** Proactive forest planning requires spatially accurate information about forest diversity. The most cost-efficient way to obtain this information is through modeling, i.e. predicting key forest diversity measures as a function of environmental factors. Patterns of forest diversity are less well known in the boreal forest of interior Alaska than in most ecosystems of North America. In order to understand the diversity patterns of this forest, we employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables. We developed prediction maps for the current levels of tree size-class and species diversity and created maps showing the potential changes to these values under a null climate change scenario and the IPCC A1B mid-range scenario for the year 2050. The method employed here yielded good accuracy for the huge Alaskan landscape despite the exclusion of spectral reflectance data due to its transient nature. The results indicate that the geographic pattern of tree species diversity differs from the pattern of tree size-class diversity across this forest type and that future climate scenarios have different effects on tree species and tree size class diversity depending on location. The results also suggest that human impact factors had a greater impact than the ecological factors in predicting the patterns of diversity within the boreal forest of interior Alaska.

**PURPOSE**

**Forest types within the State of Alaska**

**STATUS**

**MAINTENANCE AND UPDATE FREQUENCY   None planned**

**SPATIAL DOMAIN**

**BOUNDING COORDINATES**

**WEST BOUNDING COORDINATE   -162.94478**

**EAST BOUNDING COORDINATE   -137.944993**

**NORTH BOUNDING COORDINATE   68.98776**

SOUTH BOUNDING COORDINATE 59.094566

#### KEYWORDS

##### THEME

THEME KEYWORD THESAURUS None

THEME KEYWORD Forest types, Landscape heterogeneity

##### PLACE

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska, Boreal forest

#### ACCESS CONSTRAINTS

None

#### USE CONSTRAINTS

None

#### NATIVE DATA SET ENVIRONMENT

Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
10.0.3.3600

### Spatial Data Organization

DIRECT SPATIAL REFERENCE METHOD Raster

#### RASTER OBJECT INFORMATION

RASTER OBJECT TYPE Pixel

ROW COUNT 8178

COLUMN COUNT 13134

### Spatial Reference

#### HORIZONTAL COORDINATE SYSTEM DEFINITION

PLANAR

#### MAP PROJECTION

MAP PROJECTION NAME NAD 1983 Alaska Albers

ALBERS CONICAL EQUAL AREA

STANDARD PARALLEL 55.0

STANDARD PARALLEL 65.0

LONGITUDE OF CENTRAL MERIDIAN

LATITUDE OF PROJECTION ORIGIN

FALSE EASTING 0.0

FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION  
 PLANAR COORDINATE ENCODING METHOD coordinate pair  
 COORDINATE REPRESENTATION  
 ABSCISSA RESOLUTION 0.0000000030536018158500163  
 ORDINATE RESOLUTION 0.0000000030536018158500163  
 PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
 HORIZONTAL DATUM NAME D North American 1983  
 ELLIPSOID NAME GRS 1980  
 SEMI-MAJOR AXIS 6378137.0  
 DENOMINATOR OF FLATTENING RATIO 298.257222101

### Entities and Attributes

DETAILED DESCRIPTION  
 ENTITY TYPE  
 ENTITY TYPE LABEL Interior forest.tif.vat

ATTRIBUTE  
 ATTRIBUTE LABEL OID  
 ATTRIBUTE DEFINITION  
 Internal feature number.  
 ATTRIBUTE DEFINITION SOURCE ESRI  
 ATTRIBUTE DOMAIN VALUES  
 UNREPRESENTABLE DOMAIN  
 Sequential unique whole numbers that are automatically generated.

ATTRIBUTE  
 ATTRIBUTE LABEL VALUE  
 ATTRIBUTE DEFINITION  
 122 White Spruce 125 Black Spruce 901 Aspen 902 Paper Birch 904 Balsam  
 Poplar

ATTRIBUTE  
 ATTRIBUTE LABEL COUNT

### Metadata Reference

METADATA DATE 2012-08-11  
 METADATA CONTACT  
 CONTACT INFORMATION  
 CONTACT ORGANIZATION PRIMARY  
 CONTACT ORGANIZATION University of Alaska Fairbanks

**CONTACT PERSON** Brian Young

**CONTACT ELECTRONIC MAIL ADDRESS** bdyoung@alaska.edu

**METADATA STANDARD NAME** FGDC Content Standard for Digital  
Geospatial Metadata

**METADATA STANDARD VERSION** FGDC-STD-001-1998

**METADATA TIME CONVENTION** local time



**Appendix 3.3: Current predicted tree species diversity ( $H_s$ ) in the Alaskan boreal forest  
Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Current predicted tree species diversity ( $H_s$ ) in the Alaskan boreal forest**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Modeling and mapping forest diversity of interior Alaska at 1-km<sup>2</sup> resolution for current and possible future climate conditions

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that modeled and mapped current and possible future forest diversity patterns within the boreal forest of Alaska. The modeling effort employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables.

The region of interest for this study was the Alaskan boreal forest which extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains. The abstract for the paper in which this data was originally published is below for reference:

**Title:** Modeling and mapping forest diversity of interior Alaska at 1-km<sup>2</sup> resolution for current and possible future climate conditions

**Brian Young**<sup>1</sup>,

**John Yarie**<sup>1</sup>,

**David Verbyla**<sup>1</sup>,

**Falk Huettmann**<sup>2</sup>,

**F. Stuart Chapin III**<sup>3</sup>

1Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

2Institute of Arctic Biology, Biology & Wildlife Department – EWHALE lab  
– University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

3Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** Proactive forest planning requires spatially accurate information about forest diversity. The most cost-efficient way to obtain this information is through modeling, i.e. predicting key forest diversity measures as a function of environmental factors. Patterns of forest diversity are less well known in the boreal forest of interior Alaska than in most ecosystems of North America. In order to understand the diversity patterns of this forest, we employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables. We developed prediction maps for the current levels of tree size-class and species diversity and created maps showing the potential changes to these values under a null climate change scenario and the IPCC A1B mid-range scenario for the year 2050. The method employed here yielded good accuracy for the huge Alaskan landscape despite the exclusion of spectral reflectance data due to its transient nature. The results indicate that the geographic pattern of tree species diversity differs from the pattern of tree size-class diversity across this forest type and that future climate scenarios have different effects on tree species and tree size class diversity depending on location. The results also suggest that human impact factors had a greater impact than the ecological factors in predicting the patterns of diversity within the boreal forest of interior Alaska.

#### **PURPOSE**

Modeling of tree size and species diversity across the boreal forest of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

**KEYWORDS****THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Predictive mapping; tree species diversity; tree size-class diversity; machine learning (random forest); Alaska; climate change

**PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska, Boreal forest

**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**

Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS 10.0.3.3600

**Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

**RASTER OBJECT INFORMATION**

RASTER OBJECT TYPE Pixel

ROW COUNT 1363

COLUMN COUNT 1215

**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION****PLANAR****MAP PROJECTION**

MAP PROJECTION NAME NAD 1983 Alaska Albers

ALBERS CONICAL EQUAL AREA

STANDARD PARALLEL 55.0

STANDARD PARALLEL 65.0

LONGITUDE OF CENTRAL MERIDIAN

LATITUDE OF PROJECTION ORIGIN

FALSE EASTING 0.0

FALSE NORTHING 0.0

**PLANAR COORDINATE INFORMATION**

PLANAR COORDINATE ENCODING METHOD coordinate pair  
COORDINATE REPRESENTATION  
ABSCISSA RESOLUTION 0.0000000030536018158500163  
ORDINATE RESOLUTION 0.0000000030536018158500163  
PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-14  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu

METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time

**Appendix 3.4: Current predicted tree size-class diversity ( $H_d$ ) in the Alaskan boreal forest  
Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Current predicted tree size-class diversity ( $H_d$ ) in the Alaskan boreal forest**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Modeling and mapping forest diversity of interior Alaska at 1-km<sup>2</sup> resolution for current and possible future climate conditions

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that modeled and mapped current and possible future forest diversity patterns within the boreal forest of Alaska. The modeling effort employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables.

The region of interest for this study was the Alaskan boreal forest which extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains.

The abstract for the paper in which this data was originally published is below for reference:

Title: Modeling and mapping forest diversity of interior Alaska at 1-km<sup>2</sup> resolution for current and possible future climate conditions

Brian Young<sup>1</sup>,

John Yariel,

David Verbyla<sup>1</sup>,

Falk Huettmann<sup>2</sup>,

F. Stuart Chapin III<sup>3</sup>

<sup>1</sup>Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

<sup>2</sup>Institute of Arctic Biology, Biology & Wildlife Department – EWHALE lab – University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

<sup>3</sup>Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** Proactive forest planning requires spatially accurate information about forest diversity. The most cost-efficient way to obtain this information is through modeling, i.e. predicting key forest diversity measures as a function of environmental factors. Patterns of forest diversity are less well known in the boreal forest of interior Alaska than in most ecosystems of North America. In order to understand the diversity patterns of this forest, we employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables. We developed prediction maps for the current levels of tree size-class and species diversity and created maps showing the potential changes to these values under a null climate change scenario and the IPCC A1B mid-range scenario for the year 2050. The method employed here yielded good accuracy for the huge Alaskan landscape despite the exclusion of spectral reflectance data due to its transient nature. The results indicate that the geographic pattern of tree species diversity differs from the pattern of tree size-class diversity across this forest type and that future climate scenarios have different effects on tree species and tree size class diversity depending on location. The results also suggest that human impact factors had a greater impact than the ecological factors in predicting the patterns of diversity within the boreal forest of interior Alaska.

#### **PURPOSE**

Modeling of tree size and species diversity across the boreal forest of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

**KEYWORDS**

**THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Predictive mapping; tree species diversity; tree size-class diversity; machine learning (random forest); Alaska; climate change

**PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska, Boreal forest

**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**

Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS 10.0.3.3600

**Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

**RASTER OBJECT INFORMATION**

RASTER OBJECT TYPE Pixel

ROW COUNT 1363

COLUMN COUNT 1215

**Spatial Reference**

**HORIZONTAL COORDINATE SYSTEM DEFINITION**

PLANAR

**MAP PROJECTION**

MAP PROJECTION NAME NAD 1983 Alaska Albers

ALBERS CONICAL EQUAL AREA

STANDARD PARALLEL 55.0

STANDARD PARALLEL 65.0

LONGITUDE OF CENTRAL MERIDIAN

LATITUDE OF PROJECTION ORIGIN

FALSE EASTING 0.0

FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION  
PLANAR COORDINATE ENCODING METHOD coordinate pair  
COORDINATE REPRESENTATION  
ABSCISSA RESOLUTION 0.0000000030536018158500163  
ORDINATE RESOLUTION 0.0000000030536018158500163  
PLANAR DISTANCE UNITS Meter

GEODETTIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-14  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu

METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time



**Appendix 3.5: Predicted changes in tree species diversity ( $H_s$ ) from its current level to the year 2050 for a null climate change scenario Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Predicted changes in tree species diversity ( $H_s$ ) from its current level to the year 2050 for a null climate change scenario**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that modeled and mapped current and possible future forest diversity patterns within the boreal forest of Alaska. The modeling effort employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables.

The region of interest for this study was the Alaskan boreal forest which extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains.

The abstract for the paper in which this data was originally published is below for reference:

**Title:** Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions

**Brian Young**<sup>1</sup>,

**John Yariel**,

David Verbyla<sup>1</sup>,

Falk Huettmann<sup>2</sup>,

F. Stuart Chapin III<sup>3</sup>

<sup>1</sup>Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

<sup>2</sup>Institute of Arctic Biology, Biology & Wildlife Department – EWHALE lab – University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

<sup>3</sup>Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** Proactive forest planning requires spatially accurate information about forest diversity. The most cost-efficient way to obtain this information is through modeling, i.e. predicting key forest diversity measures as a function of environmental factors. Patterns of forest diversity are less well known in the boreal forest of interior Alaska than in most ecosystems of North America. In order to understand the diversity patterns of this forest, we employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables. We developed prediction maps for the current levels of tree size-class and species diversity and created maps showing the potential changes to these values under a null climate change scenario and the IPCC A1B mid-range scenario for the year 2050. The method employed here yielded good accuracy for the huge Alaskan landscape despite the exclusion of spectral reflectance data due to its transient nature. The results indicate that the geographic pattern of tree species diversity differs from the pattern of tree size-class diversity across this forest type and that future climate scenarios have different effects on tree species and tree size class diversity depending on location. The results also suggest that human impact factors had a greater impact than the ecological factors in predicting the patterns of diversity within the boreal forest of interior Alaska.

#### **PURPOSE**

Modeling of tree size and species diversity across the boreal forest of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY   None planned

#### **SPATIAL DOMAIN**

**BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478  
EAST BOUNDING COORDINATE -137.944993  
NORTH BOUNDING COORDINATE 68.98776  
SOUTH BOUNDING COORDINATE 59.094566

**KEYWORDS****THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Predictive mapping; tree species diversity; tree size-class diversity; machine learning (random forest); Alaska; climate change

**PLACE**

PLACE KEYWORD THESAURUS None

PLACE KEYWORD Alaska, Boreal forest

**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**

Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS 10.0.3.3600

**Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

**RASTER OBJECT INFORMATION**

RASTER OBJECT TYPE Pixel

ROW COUNT 1363

COLUMN COUNT 1215

**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION**

PLANAR

**MAP PROJECTION**

MAP PROJECTION NAME NAD 1983 Alaska Albers

ALBERS CONICAL EQUAL AREA

STANDARD PARALLEL 55.0

STANDARD PARALLEL 65.0

FALSE EASTING 0.0  
FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION  
PLANAR COORDINATE ENCODING METHOD coordinate pair  
COORDINATE REPRESENTATION  
ABSCISSA RESOLUTION 0.0000000030536018158500163  
ORDINATE RESOLUTION 0.0000000030536018158500163  
PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-14  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu

METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time

**Appendix 3.6: Predicted changes in tree species diversity ( $H_s$ ) from its current level to the year 2050 for a IPCC A1B scenario Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Predicted changes in tree species diversity ( $H_s$ ) from its current level to the year 2050 for a IPCC A1B scenario**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that modeled and mapped current and possible future forest diversity patterns within the boreal forest of Alaska. The modeling effort employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables.

The region of interest for this study was the Alaskan boreal forest which extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains.

The abstract for the paper in which this data was originally published is below for reference:

**Title: Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions**

**Brian Young**<sup>1</sup>,

**John Yarie**<sup>1</sup>,

David Verbyla<sup>1</sup>,

Falk Huettmann<sup>2</sup>,

F. Stuart Chapin III<sup>3</sup>

<sup>1</sup>Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

<sup>2</sup>Institute of Arctic Biology, Biology & Wildlife Department – EWHALE lab – University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

<sup>3</sup>Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** Proactive forest planning requires spatially accurate information about forest diversity. The most cost-efficient way to obtain this information is through modeling, i.e. predicting key forest diversity measures as a function of environmental factors. Patterns of forest diversity are less well known in the boreal forest of interior Alaska than in most ecosystems of North America. In order to understand the diversity patterns of this forest, we employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables. We developed prediction maps for the current levels of tree size-class and species diversity and created maps showing the potential changes to these values under a null climate change scenario and the IPCC A1B mid-range scenario for the year 2050. The method employed here yielded good accuracy for the huge Alaskan landscape despite the exclusion of spectral reflectance data due to its transient nature. The results indicate that the geographic pattern of tree species diversity differs from the pattern of tree size-class diversity across this forest type and that future climate scenarios have different effects on tree species and tree size class diversity depending on location. The results also suggest that human impact factors had a greater impact than the ecological factors in predicting the patterns of diversity within the boreal forest of interior Alaska.

#### **PURPOSE**

Modeling of tree size and species diversity across the boreal forest of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

**BOUNDING COORDINATES****WEST BOUNDING COORDINATE -162.94478****EAST BOUNDING COORDINATE -137.944993****NORTH BOUNDING COORDINATE 68.98776****SOUTH BOUNDING COORDINATE 59.094566****KEYWORDS****THEME****THEME KEYWORD THESAURUS None****THEME KEYWORD Predictive mapping; tree species diversity; tree size-class diversity; machine learning (random forest); Alaska; climate change****PLACE****PLACE KEYWORD THESAURUS None****PLACE KEYWORD Alaska, Boreal forest****ACCESS CONSTRAINTS****None****USE CONSTRAINTS****None****NATIVE DATA SET ENVIRONMENT****Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS 10.0.3.3600****Spatial Data Organization****DIRECT SPATIAL REFERENCE METHOD Raster****RASTER OBJECT INFORMATION****RASTER OBJECT TYPE Pixel****ROW COUNT 1363****COLUMN COUNT 1215****Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION****PLANAR****MAP PROJECTION****MAP PROJECTION NAME NAD 1983 Alaska Albers****ALBERS CONICAL EQUAL AREA****STANDARD PARALLEL 55.0****STANDARD PARALLEL 65.0**

LONGITUDE OF CENTRAL MERIDIAN  
LATITUDE OF PROJECTION ORIGIN  
FALSE EASTING 0.0  
FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION  
PLANAR COORDINATE ENCODING METHOD coordinate pair  
COORDINATE REPRESENTATION  
ABSCISSA RESOLUTION 0.0000000030536018158500163  
ORDINATE RESOLUTION 0.0000000030536018158500163  
PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

#### **Metadata Reference**

METADATA DATE 2012-08-14  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu

METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time



**Appendix 3.7: Predicted changes in tree size-class diversity ( $H_d$ ) from its current level to the year 2050 for a null climate change scenario Raster Dataset Metadata Standard  
Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Predicted changes in tree size-class diversity ( $H_d$ ) from its current level to the year 2050 for a null climate change scenario**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that modeled and mapped current and possible future forest diversity patterns within the boreal forest of Alaska. The modeling effort employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables.

The region of interest for this study was the Alaskan boreal forest which extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains.

The abstract for the paper in which this data was originally published is below for reference:

**Title: Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions**

**Brian Young**<sup>1</sup>,

**John Yarie**<sup>1</sup>,

David Verbyl<sup>1</sup>,

Falk Huettmann<sup>2</sup>,

F. Stuart Chapin III<sup>3</sup>

<sup>1</sup>Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

<sup>2</sup>Institute of Arctic Biology, Biology & Wildlife Department – EWHALE lab – University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

<sup>3</sup>Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** Proactive forest planning requires spatially accurate information about forest diversity. The most cost-efficient way to obtain this information is through modeling, i.e. predicting key forest diversity measures as a function of environmental factors. Patterns of forest diversity are less well known in the boreal forest of interior Alaska than in most ecosystems of North America. In order to understand the diversity patterns of this forest, we employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables. We developed prediction maps for the current levels of tree size-class and species diversity and created maps showing the potential changes to these values under a null climate change scenario and the IPCC A1B mid-range scenario for the year 2050. The method employed here yielded good accuracy for the huge Alaskan landscape despite the exclusion of spectral reflectance data due to its transient nature. The results indicate that the geographic pattern of tree species diversity differs from the pattern of tree size-class diversity across this forest type and that future climate scenarios have different effects on tree species and tree size class diversity depending on location. The results also suggest that human impact factors had a greater impact than the ecological factors in predicting the patterns of diversity within the boreal forest of interior Alaska.

#### **PURPOSE**

Modeling of tree size and species diversity across the boreal forest of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

**BOUNDING COORDINATES****WEST BOUNDING COORDINATE** -162.94478**EAST BOUNDING COORDINATE** -137.944993**NORTH BOUNDING COORDINATE** 68.98776**SOUTH BOUNDING COORDINATE** 59.094566**KEYWORDS****THEME****THEME KEYWORD THESAURUS** None**THEME KEYWORD** Predictive mapping; tree species diversity; tree size-class diversity; machine learning (random forest); Alaska; climate change**PLACE****PLACE KEYWORD THESAURUS** None**PLACE KEYWORD** Alaska, Boreal forest**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**

Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS 10.0.3.3600

**Spatial Data Organization****DIRECT SPATIAL REFERENCE METHOD** Raster**RASTER OBJECT INFORMATION****RASTER OBJECT TYPE** Pixel**ROW COUNT** 1363**COLUMN COUNT** 1215**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION****PLANAR****MAP PROJECTION****MAP PROJECTION NAME** NAD 1983 Alaska Albers**ALBERS CONICAL EQUAL AREA****STANDARD PARALLEL** 55.0**STANDARD PARALLEL** 65.0

LONGITUDE OF CENTRAL MERIDIAN  
LATITUDE OF PROJECTION ORIGIN  
FALSE EASTING 0.0  
FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION  
PLANAR COORDINATE ENCODING METHOD coordinate pair  
COORDINATE REPRESENTATION  
ABSCISSA RESOLUTION 0.0000000030536018158500163  
ORDINATE RESOLUTION 0.0000000030536018158500163  
PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

#### **Metadata Reference**

METADATA DATE 2012-08-14  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu

METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time

**Appendix 3.8: Predicted changes in tree size-class diversity ( $H_d$ ) from its current level to the year 2050 for a IPCC A1B scenario Raster Dataset Metadata Standard Version FGDC-STD-001-1998**

**Identification**

**CITATION**

**CITATION INFORMATION**

**PUBLICATION DATE** 2011-04-06

**PUBLICATION TIME** 000000

**TITLE**

**Predicted changes in tree size-class diversity ( $H_d$ ) from its current level to the year 2050 for a IPCC A1B scenario**

**GEOSPATIAL DATA PRESENTATION FORM** raster digital data

**SERIES INFORMATION**

**SERIES NAME** Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions

**DESCRIPTION**

**ABSTRACT**

This raster layer is from a study that modeled and mapped current and possible future forest diversity patterns within the boreal forest of Alaska. The modeling effort employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables.

The region of interest for this study was the Alaskan boreal forest which extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains.

The abstract for the paper in which this data was originally published is below for reference:

**Title: Modeling and mapping forest diversity of interior Alaska at 1-km2 resolution for current and possible future climate conditions**

**Brian Young**<sup>1</sup>,

**John Yarie**<sup>1</sup>,

David Verbyla<sup>1</sup>,

Falk Huettmann<sup>2</sup>,

F. Stuart Chapin III<sup>3</sup>

<sup>1</sup>Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

<sup>2</sup>Institute of Arctic Biology, Biology & Wildlife Department – EWHALE lab – University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

<sup>3</sup>Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** Proactive forest planning requires spatially accurate information about forest diversity. The most cost-efficient way to obtain this information is through modeling, i.e. predicting key forest diversity measures as a function of environmental factors. Patterns of forest diversity are less well known in the boreal forest of interior Alaska than in most ecosystems of North America. In order to understand the diversity patterns of this forest, we employed Random Forest analysis (machine learning) and the Boruta algorithm to predict tree species and tree size-class diversity for the entire region using a combination of forest inventory data and a suite of 28 predictors from public open-access data archives that included climatic, soil, distance, and topographic variables. We developed prediction maps for the current levels of tree size-class and species diversity and created maps showing the potential changes to these values under a null climate change scenario and the IPCC A1B mid-range scenario for the year 2050. The method employed here yielded good accuracy for the huge Alaskan landscape despite the exclusion of spectral reflectance data due to its transient nature. The results indicate that the geographic pattern of tree species diversity differs from the pattern of tree size-class diversity across this forest type and that future climate scenarios have different effects on tree species and tree size class diversity depending on location. The results also suggest that human impact factors had a greater impact than the ecological factors in predicting the patterns of diversity within the boreal forest of interior Alaska.

#### **PURPOSE**

Modeling of tree size and species diversity across the boreal forest of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

**BOUNDING COORDINATES****WEST BOUNDING COORDINATE** -162.94478**EAST BOUNDING COORDINATE** -137.944993**NORTH BOUNDING COORDINATE** 68.98776**SOUTH BOUNDING COORDINATE** 59.094566**KEYWORDS****THEME****THEME KEYWORD THESAURUS** None**THEME KEYWORD** Predictive mapping; tree species diversity; tree size-class diversity; machine learning (random forest); Alaska; climate change**PLACE****PLACE KEYWORD THESAURUS** None**PLACE KEYWORD** Alaska, Boreal forest**ACCESS CONSTRAINTS**

None

**USE CONSTRAINTS**

None

**NATIVE DATA SET ENVIRONMENT**

Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS 10.0.3.3600

**Spatial Data Organization****DIRECT SPATIAL REFERENCE METHOD** Raster**RASTER OBJECT INFORMATION****RASTER OBJECT TYPE** Pixel**ROW COUNT** 1363**COLUMN COUNT** 1215**Spatial Reference****HORIZONTAL COORDINATE SYSTEM DEFINITION****PLANAR****MAP PROJECTION****MAP PROJECTION NAME** NAD 1983 Alaska Albers**ALBERS CONICAL EQUAL AREA****STANDARD PARALLEL** 55.0**STANDARD PARALLEL** 65.0

LONGITUDE OF CENTRAL MERIDIAN  
 LATITUDE OF PROJECTION ORIGIN  
 FALSE EASTING 0.0  
 FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION  
 PLANAR COORDINATE ENCODING METHOD coordinate pair  
 COORDINATE REPRESENTATION  
 ABSCISSA RESOLUTION 0.0000000030536018158500163  
 ORDINATE RESOLUTION 0.0000000030536018158500163  
 PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
 HORIZONTAL DATUM NAME D North American 1983  
 ELLIPSOID NAME GRS 1980  
 SEMI-MAJOR AXIS 6378137.0  
 DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-14  
 METADATA CONTACT  
 CONTACT INFORMATION  
 CONTACT ORGANIZATION PRIMARY  
 CONTACT ORGANIZATION University of Alaska Fairbanks  
 CONTACT PERSON Brian Young  
 CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu

METADATA STANDARD NAME FGDC Content Standard for Digital  
 Geospatial Metadata  
 METADATA STANDARD VERSION FGDC-STD-001-1998  
 METADATA TIME CONVENTION local time



## CHAPTER 4: MAPPING AND PREDICTING ABOVEGROUND BIOMASS OF TREES USING FOREST INVENTORY DATA AND PUBLIC ENVIRONMENTAL VARIABLES WITHIN THE ALASKAN BOREAL FOREST<sup>4.1</sup>

### 4.1 ABSTRACT

A method for mapping the predicted forest biomass was developed and tested on a study region in the boreal forest of interior Alaska. In order to understand aboveground biomass values within this forest, we employed the Boruta Algorithm, Random Forest analysis, and Regression Tree analysis (three different machine learning techniques), as well as tests for spatial autocorrelation, to predict aboveground woody biomass for the entire region using a combination of forest inventory data and a suite of 32 predictors from public open-access data archives that included spectral reflectance, climatic, soil, distance from various features, and topographic variables. We also developed, for the first time, high-resolution prediction maps at a 1km<sup>2</sup> pixel size for aboveground woody biomass. The method employed here yielded good accuracy for the huge Alaskan landscape despite a rather limited training set. The results indicate that the geographic patterns of biomass are strongly influenced by the tree size-class diversity of a given stand.

### 4.2 INTRODUCTION

Forest biomass, the aboveground dry mass portion of live trees within a given area (Bonnor 1985), is of particular interest for both ecological and economic reasons. Forest soils and forest biomass hold most of the carbon in the Earth's terrestrial biomes (Houghton 2005) and contribute significantly to the global carbon cycle (Schimel et al.

---

<sup>4.1</sup> Prepared for submission to Western Journal of Applied Forestry as: Young B, Yarie J, Verbyla D, Huettmann F, Chapin FS. Mapping and predicting aboveground biomass of trees using forest inventory data and public environmental variables within the Alaskan boreal forest

2001). Within the boreal forest, one of the largest terrestrial biomes, little is known about the quantity of woody biomass at spatial scales useful to forest practitioners. The scale at which forestry typically operates is at the intermediate  $\sim 1\text{m}^2$  to  $3.0\text{km}^2$ , or mesoscale (Niemelä 1999; O'Neill et al. 1997). Previous investigations of aboveground forest biomass in the boreal region (see for instance Blackard et al. 2008; Botkin and Simpson 1990; Harrell et al. 1995; Yarie and Billings 2002; Yarie and Mead 1982), the spatial scales have either been rather coarse (Botkin and Simpson 1990; Yarie and Billings 2002) or the predictions lack precision due to limited ground-truthing. The increased interest in biomass as a possible fuel source has raised interest among many communities in the quantities of biomass that are available to them at the operational or meso scale (Fresco 2006; GAO 2005; Loeffler et al. 2010). This subject represents a natural resource topic and is of national, state, economical, ecological, and strategic interest.

The combination of forest inventory data with remote sensing data from both aerial and satellite formats have been previously employed in mapping woody biomes across broad spatial scales (see for instance Fassnacht et al. 2006; Franklin 2001; Iverson and Prasad 2001; McRoberts et al. 2008; Ruefenacht et al. 2008). In the United States, the inventory data is typically from the Forest Inventory and Analysis (FIA) program of the USDA Forest Service which is uniformly distributed across the landscape, except in interior Alaska, which has not had even a partial inventory in over 25 years. Spatial interpolation techniques that combine forest-inventory and remote-sensing data in regions with either sparse or non-uniformly distributed inventory data have been employed to predict the geographical distribution of various forest attributes (Liang and Zhou 2010; Parmentier et al. 2011; Young et al. 2011). The use of machine learning to estimate aboveground forest biomass at the mesoscale from forest inventory and remotely sensed data in remote boreal forest regions has, however, received less attention.

Maps depicting spatially explicit estimates of forest biomass are valuable for planning and monitoring (Drew et al. 2010). Creating such maps typically employs predictive spatial modeling techniques where the parameters of interest are obtained from inventory data and then related to remotely mapped attributes (see Austin 2002; Cushman

and Huettmann 2010; Cushman and McKelvey 2009; Ferrier et al. 2002; Franklin 1995; Guisan and Thuiller 2005). Numerous statistical approaches have been used to create such maps, with non-parametric approaches tending to yield better results than parametric approaches, which often violate required statistical assumptions (Drew et al. 2010; Prasad et al. 2006). The use of machine learning, and notably the random forest algorithm, has allowed for major advances in the capacity to make predictions of various forest attributes including biomass (Baccini et al. 2008; Craig and Huettmann 2008; Li et al. 2011). Furthermore, addressing spatial autocorrelation (Bivand et al. 2008; Legendre 1993), especially in large-scale forest studies (Liang and Zhou 2010), is crucial because when unaccounted for it may affect statistical model predictions due to violation of independence on which most standard statistical procedures rely (Legendre 1993). Thus, non-parametric models that account for, or are at least tolerant of, spatial autocorrelation and noisy data could be generally useful in assessing the spatial patterns of aboveground biomass (Craig and Huettmann 2008; Li et al. 2011).

The boreal forest of Alaska extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Chugach and Coastal mountains (Figure 4.1), covering an area of nearly 500,000 km<sup>2</sup>. The Alaskan boreal forest consists of a mosaic of two general forest types, mixed aspen/birch and mixed spruce (Ruefenacht et al. 2008; Viereck and Little 2007; Young et al. 2011) and primarily contains seven tree species. White spruce (*Picea glauca*) and black spruce (*P. marianana*) are the predominant conifers and two poplars (*Populus tremuloides* and *P. balsamifera*), two birches (*Betula neoalaska* and *B. kенаica*), and tamarack (*Larix laricina*) represent the deciduous species. Significant variation in tree growth occurs due to local differences in topography, soil type, biota, successional state, and climate conditions (Chapin et al. 2006; Liang 2010; Lloyd and Fastie 2002; Van Cleve et al. 1983; Wilmking and Juday 2005), which are collectively referred to as state factors (Major 1951). These variations in tree growth can lead to vastly different amounts of aboveground forest biomass depending on site differences in state factors. Biomass models that incorporate as many of the state factors as possible

may yield an enhanced predictive ability for this ecologically and economically important forest attribute (Cutler et al. 2007; Grossmann et al. 2010).

Our objectives here are to 1) develop a spatial model depicting aboveground forest biomass for the Alaskan boreal forest from ground-measured inventory plots, and then extrapolate to a 1-km cell size, 2) evaluate model performance, 3) explore the contribution of the many environmental predictors used to develop the biomass model, and 4) develop stand-level predictions of forest biomass using remote sensing data and geographic information system (GIS) tools.

### 4.3 METHODS AND MATERIALS

#### 4.3.1 BIOMASS DATA

Our dataset consisted of 704 permanent sample plots (PSPs) from the Cooperative Alaska Forest Inventory Database (CAFI; [http://www.lter.uaf.edu/data\\_detail.cfm?datafile\\_pkey=452](http://www.lter.uaf.edu/data_detail.cfm?datafile_pkey=452)) (Malone et al. 2009) and the Fort Wainwright Forest Inventory Database (WAIN; <http://www.usarak.army.mil/conservation>) (Rees, personal communication). These databases consist of periodically re-measured PSPs located across interior and south-central Alaska north of 60°N (Fig. 1). The 429 CAFI plots are 0.04 ha in size and are primarily located on well-stocked forested areas along the road system on Federal, State, Borough, and Native Corporation lands (Malone et al. 2009) while the 265 WAIN plots are scattered across forested areas on Military lands (Figure 4.1) and were established as per the FIA protocol (Miles et al. 2001).

The aboveground tree woody biomass, which includes biomass from the tree bole, stumps, branches and twigs, on each of the 704 PSPs was calculated for each tree. This was done for trees greater than 2.54 cm in diameter at breast height (DBH) using the equations developed by Jenkins et al. (2003) for each of the seven possible tree species (*Picea glauca*, *P. marianana*, *Populus tremuloides*, *P. balsamifera*, *Betula neoalaska*, *B. kenaica*, and *Larix laricina*) present within a given PSP. These were then aggregated to

develop a megaton per hectare (Mg/ha) dry weight value. The Jenkins et al. (2003) calculations are used by the United States Forest Service Forest Inventory and Analysis Program (FIA) for their PSPs (Jenkins et al. 2004). We choose to use the same calculations even though regional biomass calculations have been developed (see Yarie et al. 2007) so that the results from our model of aboveground tree woody biomass could be directly compared with other previously published results (Blackard et al. 2008).

#### 4.3.2 ENVIRONMENTAL FACTORS

Our predictor dataset consisted of 39 variables, including the spatial structure, climatic, topographic, vegetation, anthropogenic, and geophysical variables (Table 4.1). The climate variables were obtained from the Scenarios Network for Alaska Planning (SNAP; <http://www.snap.uaf.edu/downloads/alaska-climate-datasets>) which contains historical downscaled datasets derived from Climate Research Unit (CRU) and Parameter-elevation Regressions on Independent Slopes Model (PRISM) data, which have been shown to perform well in Alaska (Walsh et al. 2008). The spatial resolution of the monthly and annual temperature data are at 2km<sup>2</sup> grid size and were averaged over the years 1901-2009 while, the monthly and annual precipitation data are averaged from 1901-2006. The topographic variables were derived from 300m digital elevation models (Alaska Geospatial Data Clearinghouse (AGDC; <http://agdc.usgs.gov/agdc.html>) using Spatial Analysis surface analysis tool and the TPI extension for ArcGis (Jenness 2006) within ArcGis 10.0 (ESRI 2011). The vegetation variable of vegetation type was derived from classifications developed using the phenology of a vegetation index (AVHRR/NDVI) collected during the 1991 growing season (Table 4.2; Fleming 1997; <http://agdc.usgs.gov/data/projects/fhm/#G>). Forest type was obtained from the 250m Forest Type Groups of Alaska map (Ruefenacht et al. 2008; [http://fsgeodata.fs.fed.us/rastergateway/forest\\_type/](http://fsgeodata.fs.fed.us/rastergateway/forest_type/)) for the year 2008. The NDVI (normalized difference vegetation index) values were obtained from the 14-day band 6 Advanced Very High Resolution Radiometer (AVHRR) data from the NOAA polar-orbiting satellites covering the periods May through September 2011, which were

obtained from the Geographic Information Network of Alaska database (GINA; [http://docs.gina.alaska.edu/ndvi/how\\_to.html](http://docs.gina.alaska.edu/ndvi/how_to.html)). The maximum NDVI variable (Gmax) was created using the maximum value from this time period while the mean NDVI growing season variable (Gmean) used the mean value from this same period of time. The stand age was determined using data from the Alaska Interagency Coordination Center ([http://afsmaps.blm.gov/imf\\_fire/imf.jsp?site=fire](http://afsmaps.blm.gov/imf_fire/imf.jsp?site=fire)) fire perimeter data covering the years 1942 – 2011. Using this fire history data, a binary response variable was created for each location describing the forests as young (< 69 years) or mature (> 69 years) based on when the location last burned (Johnson et al. 2011). The tree size basal area and tree species basal area diversity values were obtained from Young et al. (Chapter 3). We calculated the anthropogenic and the geophysical variables using data from AGDC and tools within ArcGis 10.0 (ESRI 2011). The predictor variables with a spatial resolution greater than 1km underwent either nearest neighbor resampling, if the data were categorical, or bilinear interpolation resampling, if the data were continuous. Those variables that had a native spatial resolution smaller than 1km<sup>2</sup> were rescaled. The predictor dataset was constructed by overlaying the individual datasets in ArcGis 10.0, then at each PSP location the environmental variables were extracted, resulting in a table with aboveground tree woody biomass values as the response variable and the environmental variables as predictors (as per Ohse et al., 2009).

#### 4.3.3 THE CALIBRATION AND VALIDATION DATASETS

For the 704 sites within our study region we had information about the aboveground forest woody biomass and the 39 environmental factors plus the latitude and longitude for each of the PSPs (Table 4.1). The dataset was randomly split into a calibration dataset (Cal,  $n = 528$ ; 75% of the plots) and a validation dataset (Val,  $n = 176$ ; 25% of the plots). The relationship between aboveground forest woody biomass and the environmental factors was modeled using the calibration dataset and the quality of the predictions was assessed using the validation dataset.

#### 4.3.4 STATISTICAL METHODS

In order to model aboveground forest biomass we determined the association between biomass and the environmental predictors at each of the PSP locations by using a two-step approach. We first used random forest analysis (RFA; Breiman 2001) in the randomForest package (Liaw and Wiener 2002) for the R system, in order to estimate and rank the importance of the predictors and to examine the nature of the relationships between the response variables and the important predictor variables (Breiman 2001). We then employed regression tree analysis (RTA; De'ath and Fabricius 2000) in the Party package (Hothorn et al. 2006a) for the R system, to further examine the relationships between aboveground forest biomass and its important environmental predictors. The specific type of RTA employed here was conditional inference-based regression trees (Hothorn et al. 2006b), a significant improvement over previous classification and regression tree algorithms (e.g. Breiman et al. 1984), which implement unbiased variable selection based on permutation tests and stopping rules (Hothorn et al. 2006b). These two non-parametric methods (RFA and RTA) are ideally suited for analyzing noisy environmental data and are able to take into account the complex and known ecological and environmental interactions between the variables (Cutler et al. 2007; De'ath and Fabricius 2000; Prasad et al. 2006).

Prior to implementing the RFA or the RTA, we used the Boruta algorithm (Kursa and Rudnicki 2010) implemented in the R system, to select the significant environmental predictors for aboveground forest biomass. This algorithm adds several random permutations to each of the original variables (Table 4.1) by running Random Forest algorithms several times (in this case 500). The statistical significance for each variable was then determined by comparing the original variable value to the value created via random permutation using Z-scores (in this study  $|Z| \leq 1.65$ ) (Kursa and Rudnicki 2010).

The final set of variables that were selected using the Boruta algorithm was then used in the RFA on the calibration (Cal) dataset. RFA models are built through a multi-step process (for details, see Breiman 2001). First, a bootstrap sample was selected from the plot data and, for these data; a regression tree, a type of prediction model that can be

represented as a decision tree, was built from the sample. Each node within the tree was constructed by selecting a random subset of the environmental variables and then determining which variable yielded the most effective split for maximizing the purity in the two resultant groups. Additional nodes were then continuously added to the tree until there was only one plot per resultant leaf. This process was repeated until the desired number of trees was built (2,500 in this analysis). To obtain a prediction from the forest of regression trees, an average of all the trees was then taken. The importance value for each predictor was then calculated by investigating the percent increase in mean squared error (MSE).

To evaluate the results of the final RFA aboveground forest biomass model, we assessed the model's predictive ability by testing it on an independent validation dataset (the Val data set with 176 measures of biomass and environmental factors). In order to validate the results of our RFA model we used root mean square error (*RMSE*), mean absolute error (*MAE*), and Pearson's product-moment correlation coefficients (*r*) on both the calibration (Cal) and the validation (Val) datasets as follows:

$$RMSE = \left[ \frac{1}{n} \sum_{i=1}^n (p_i - o_i)^2 \right]^{\frac{1}{2}} \quad (4.1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_i - o_i| \quad (4.2)$$

$$r = \frac{\sum_{i=1}^n (p_i - \bar{p})(o_i - \bar{o})}{\sqrt{\sum_{i=1}^n (p_i - \bar{p})^2 \sum_{i=1}^n (o_i - \bar{o})^2}} \quad (4.3)$$

where n: number of plots within either the calibration (Cal) or validation (Val) datasets, o: observed value, p: predicted value,  $\bar{o}$ : mean of observed values, and  $\bar{p}$  is the mean of the predicted values. In addition, to provide a way to visualize the marginal effects of the predictor variables in the RFA estimates of aboveground forest biomass we developed partial dependence plots using the randomForest package in the R system.



Following the RFA analysis we employed RTA to enhance our understanding of the relationships between the six most important environmental predictors for aboveground forest woody biomass determined through the RFA process. RTA recursively partitions a dataset into subsets that are relatively homogeneous with regards to the response variable (De'ath and Fabricius 2000). The RTA as applied to these data used conditional inference trees (Hothorn et al. 2006b), which require a statistical significant P value ( $P < 0.001$  for this analysis) determined through a Monte Carlo randomization procedure (9999 permutations used in this analysis). This technique minimizes selection bias and over-fitting, common problem in recursive binary partitioning methods, by using a combination of pruning procedures and a stopping criteria (Hothorn et al. 2006b). The results of the RTA are easily interpreted as a classification tree containing a set of decision rules on the environmental predictor variables.

Measures of biomass, along with most environmental variables, are typically spatially autocorrelated (Bivand et al. 2008; Cressie 1993; Lamsal et al. 2012) and when untreated, violate the assumption of independence within statistical models (Li et al. 2011; Miller et al. 2007). We tested for spatial autocorrelation, as well as for large-scale spatial patterns, within the residuals of the final RFA aboveground forest biomass model. In order to conduct tests for autocorrelation, we first assumed that plots at distant locations will affect each other less than plots that are close to one another (Cressie 1993; Drew et al. 2010). We therefore applied a spatial weight of inverse distance. Given the neighborhood structure, we were then able to evaluate the residuals of the RFA model for aboveground forest biomass using Moran's I and Geary's C test statistics (see Cressie 1993 for discussion) with the *spdep* (Bivand et al. 2007) package for the R system.

#### 4.3.5 PREDICTIVE MAPS

Once calibrated and validated, the final RFA model (Table 4.3) was then applied to the entire boreal forest region of Alaska at a  $1\text{km}^2$  resolution in order to obtain an estimate of the aboveground forest biomass (Mg/ha dry weight). The accuracy and

confidence of this prediction was also evaluated against previously published results for this study region.

### 4.3. RESULTS

#### 4.3.1 VARIABLE SELECTION AND IMPORTANCE

The Boruta algorithm was used as the basis of our variable selection processes to predict aboveground forest biomass (biomass). The results of this process reduced the total number of variables (Table 4.1), based on z-scores, by 7 variables to the reduced set (found in Table 4.3) that were then applied as the final random forest model (RFA) to predict biomass.

Table 4.3 also shows the ranking of the predictor variables by their importance as determined by the percent increase in mean standard error (%IncMSE). The variable of tree size-class diversity ( $H_d$ ) is by far the most important variable to predict biomass for the Alaskan boreal forest accounting for over 90%IncMSE. An additional vegetation variable, Vegetation type (Veg), was the second most important variable accounting for 36.40%IncMSE. Of the climatic variables, June precipitation (P\_06) was deemed the most important accounting for over 31%IncMSE. The anthropogenic variable of Distance to communities (DTC) was also found to be highly influential in predicting biomass contributing to nearly 30%IncMSE. Of the top ten variables to predict biomass three were vegetation variables ( $H_d$ , Veg, and Maximum NDVI (maxAV)), three are topographic variables (elevation (EL); the proxy of site productivity (Prod), which was derived from elevation, slope, and aspect (Stage and Salas 2007); and slope (SI)), two climatic variables (P\_06, and winter precipitation (P\_W)), and one spatial structure variable (longitude (Y)), which is an indicator of continentality.

#### 4.3.2 RANDOM FOREST ANALYSIS MODEL ASSESSMENT

The RFA model for aboveground forest biomass (determined by using the variables within Table 4.3) as applied to the calibration dataset (Cal) was able to explain

35.23% of the total variation in biomass (Figure 4.2:  $RMSE_{cal} = 19.73$ ,  $MAE_{cal} = 13.76$ ,  $r_{cal} = 0.96$ ,  $P < 0.001$ ). The RFA model also provided good predictions on the validation dataset (Figure 4.2:  $RMSE_{val} = 43.84$ ,  $MAE_{val} = 32.14$ ,  $r_{val} = 0.55$ ,  $P < 0.001$ ). Despite these high significance levels, the RFA model for biomass tended to slightly overestimate on the validation dataset for low biomass sites and underestimate in high sites.

Additionally, we observed some variation in the magnitude of the errors; large errors are unlikely due to the relatively small difference between the  $RMSE_{val}$  and  $MAE_{val}$  values. Our modeled average differences on the validation dataset between the predicted and the observed biomass value was 32.14 Mg/ha with a mean predicted value of 92.68 Mg/ha.

In comparison to our results, the predictions of Blackard et al. (2008) for aboveground forest woody biomass on the validation dataset (Val) were noticeably different (Figure 4.3). Using the results from the model by Blackard et al. (2008) for biomass, we found a correlation coefficient of - 0.07 for the 176 PSPs. Additionally, the RMSE was 76.18 and the MAE was 59.00, suggesting a large magnitude of error and low accuracy given that the average difference is found to be greater than the mean predicted value (44.46 Mg/ha) from the Blackard et al. (2008) model for aboveground forest biomass within this study region.

#### 4.3.3 INFLUENCE OF THE ENVIRONMENTAL FACTORS ON ABOVEGROUND FOREST BIOMASS

The partial dependency plots (Figure 4.4) illustrate the relationships between aboveground forest biomass and the six most important environmental factors as determined by RFA. The most important factor, tree size-class diversity ( $H_d$ ) indicates a potential threshold at an  $H_d$  value of  $> 1.5$  biomass steeply increases but below this level, it is stable at  $\sim 70$  Mg/ha. Vegetation type (Veg) is the second most influential with closed forest vegetation types having the greatest biomass values. June precipitation ( $P_{06}$ ) also appears to express a threshold effect on biomass, with values less than 55 mm having higher biomass than those receiving more moisture. Distance to communities (Dtc) also displays a pronounced effect on biomass with decreasing biomass values with

increasing distance. The effect of Elevation (El) on aboveground woody biomass suggests a depressed value at mid-elevations with increased amounts on either extreme. The effect of continentality, as somewhat measured by Longitude (Y), is also pronounced with a slightly positive effect on biomass.

#### 4.3.4 REGRESSION TREE ANALYSIS MODEL ASSESSMENT

The RTA of aboveground forest biomass using the six most important variables determined from RFA produced 8 terminal nodes (Figure 4.5). The highest biomass values appear to occur when tree size class diversity ( $H_d$ ) is greater than 2.221 and on sites with  $H_d$  values between 1.903 and 2.221 that receive less than or equal to 54 mm of precipitation in June. The lowest biomass values occur on low diversity ( $\leq 1.583$ ) sites at elevations less than 417 m and greater than 30.7km away from a community. Additionally, low biomass values are predicted for sites greater than 417m in elevation which occupy micro sites dominated by open and or closed spruce forests or mixture of spruce shrub woodlands.

#### 4.3.5 SPATIAL DEPENDENCY OF ABOVEGROUND FOREST BIOMASS

Biomass is strongly spatially autocorrelated (Table 4.4). By incorporating numerous environmental predictor variables in the final RFA biomass model, the spatial autocorrelation present in aboveground biomass were effectively eliminated as evident by the lack of autocorrelation present in the residuals of the final model (Table 4.4).

#### 4.3.6 PREDICTED ABOVEGROUND FOREST BIOMASS PATTERNS

The aboveground forest woody biomass for the boreal forest of Alaska as predicted through Random Forest analysis using environmental predictors and forest inventory data that were then predicted across the study region ranged widely (Figure 4.6). The predicted values ranged from less than 58 Mg/ha to greater than 120 Mg/ha with a mean value of 90 Mg/ha. The highest values were primarily found within the

central interior between Fairbanks and McGrath, around Glennallen, and north of Anchorage in the Matanuska-Susitna valleys.

#### 4.4. DISCUSSION

We assessed the spatial variation of aboveground biomass across the heterogeneous landscape of the Alaskan boreal forest and developed a model to predict and map biomass at un-sampled locations. Tree size class diversity ( $H_d$ ) was more important than environmental variables in predicting aboveground forest biomass in this study. This result was not surprising because the structure of a young forest is typically characterized by a single canopy layer, high stem density, few forest gaps, and trees of roughly the same size with generally lower biomass (Pretzsch 2005; Scherer-Lorenzen et al. 2005; Schulze et al. 2005), whereas older forests generally have a greater mixture of tree sizes in multiple canopy layers due primarily to niche differentiation (Harper et al. 2003; Kneeshaw and Gauthier 2003; McCarthy 2001), resulting in higher overall stand biomass (Kohyama 1993; Scherer-Lorenzen et al. 2005). Forest age was relatively unimportant (see table 4.3) in this analysis, probably because it was represented in this dataset by a binary response, at an age of 69 years, that may not reflect the age at which significant niche differentiation occurs within this forest type. An alternative explanation could be that the environment ultimately controlled tree size class diversity and productivity. For example, a black spruce stand on permafrost-rich soils will always display lower productivity than a similarly aged white spruce stand on a southerly aspect. Thus, tree size class diversity was higher on the productive white spruce sites, while age would be unrelated to productivity at least in this case. Regardless of the reason, the results suggest that forest managers could enhance biomass production by increasing tree size diversity in the boreal forest of Alaska if biomass production becomes a primary management goal.

The AVHRR band 6 (NDVI) data (Gmax and Gmean variables) were not as important at predicting biomass within this study as some of the other vegetation, climatic, anthropogenic, and topographic variables (see table 4.3). While reflectance

variables have previously been shown to be good predictors of biomass (Baccini et al. 2004; Baccini et al. 2008; Blackard et al. 2008; Powell et al. 2010), they were overshadowed by numerous other variables (see Table 4.3) in this study. For example, the previously derived categorical variable of vegetation proved to outperform both measures of NDVI at predicting biomass, which was likely due to it being derived from a combination of NDVI and other topographical variables. Additionally, broadleaf canopies tend to have higher NDVI values when compared to conifer or mixed canopies but, have lower overall aboveground biomass at least within the boreal forest of Alaska.

The Interior Forest of Alaska is characterized by a wide range of elevation and climate zones and these variables exert important controls on the spatial distribution of aboveground biomass. For example, average June precipitation was important for separating forests into larger and smaller timber volumes (Figures 4.4 and 4.5). However, contrary to our expectations, higher biomass values were observed at lower precipitation amounts. This is perhaps because low June precipitation coincided with warmer temperatures in more continental regions. Or perhaps, spring snowmelt resulting in soil moisture recharge may have been sufficient to meet water demands in a warmer June with little precipitation present.

The presence of broadleaf trees mixed with conifers created particular difficulties, and the model tended to underestimate biomass in areas characterized by broadleaf and conifer mixtures (Table 4.3). In this context, the use of 1 km<sup>2</sup> spatial resolution was a key challenge for this work because virtually all grid cells included multiple forest stands and mixtures of forests and shrubs. Future efforts using finer (e.g., 30 m) resolution data should help to resolve this problem but problems may still arise because it is common to find white spruce in an aspen or birch stand or even birch in a black spruce stand within the study region.

Predictive models and their performance can be sensitive to sample size and information contained within the samples. Consequently, landscape and regional scale predictive models often strike a balance between sample size and prediction performance. For example, RFA models developed with small training sets are prone to low prediction

performance (Breiman 2001). However, large sample sizes can be cost-prohibitive particularly in remote locations such as Alaska. This question of required sampling effort versus predicted accuracy and low variance is a key topic for virtually all natural resource inventories. However, the use of machine learning is drastically enhancing our ability to overcome some of these limitations.

Assessment of the spatial variation of forest biomass across landscapes is challenging but vital in order to improve regional-scale assessments. However, due to the presence of autocorrelated environmental variables at multiple spatial scales, largely due to community processes (Legendre 1993), our ability to assess their spatial variation likely depends upon the spatial arrangement of our permanent sample plots (Lamsal et al. 2012). Across the heterogeneous landscapes of boreal Alaska, we had a dense but localized sampling effort that made it possible to detect localized patterns of biomass but our predictions for more distant locations are based strictly on correlation, which proves to be a very powerful approach. Therefore, landscape- to regional-scale models of forest biomass distribution within the heterogeneous boreal forest of Alaska could benefit with a wider geographic and well-designed array of permanent sample plots to assess this vital resource of relevance to Alaska and beyond.

#### 4.6 REFERENCES

- Austin, M.P. 2002. Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol. Model.* 157(2-3):101-118.
- Baccini, A., M.A. Friedl, C.E. Woodcock, and R. Warbington. 2004. Forest biomass estimation over regional scales using multisource data. *Geophysical Research Letters* 31(10):L10501.
- Baccini, A., N. Laporte, S.J. Goetz, M. Sun, and H. Dong. 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters* 3(4):045011.
- Bivand, R.S., L. Anselin, O. Berke, A. Bernat, M. Carvalho, Y. Chun, C.F. Dormann, S. Dray, R. Halbersma, and N. Lewin-Koh. 2007. *spdep: Spatial dependence: weighting schemes, statistics and models*. R package version 0.4-9.
- Bivand, R.S., E.J. Pebesma, and V. Gómez-Rubio. 2008. *Applied spatial data analysis with R*. Springer Verlag, New York, New York, USA.
- Blackard, J., M. Finco, E. Helmer, G. Holden, M. Hoppus, D. Jacobs, A. Lister, G. Moisen, M. Nelson, and R. Riemann. 2008. Mapping U.S. forest biomass using nationwide forest inventory data and moderate resolution information. *Remote Sens. Environ.* 112(4):1658-1677.
- Bonnor, G. 1985. *Inventory of forest biomass in Canada*. Canadian Forestry Service, Environment Canada, Forestry Statistics and Systems Branch, Ontario, Canada.
- Botkin, D.B., and L.G. Simpson. 1990. Biomass of the North-American Boreal Forest - a Step toward Accurate Global Measures. *Biogeochemistry* 9(2):161-174.
- Breiman, L. 2001. Random forests. *Mach. Learn.* 45(1):5-32.



Breiman, L., J. Friedman, R. Ohlsen, and C. Stone. 1984. Classification and regression trees. Chapman and Hall/CRC Press, New York

Chapin, F.S., III, T. Hollingsworth, D.F. Murray, L.A. Viereck, and M.D. Walker. 2006. Floristic diversity and vegetation distribution in the Alaskan boreal forest. P. 81-99 in Alaska's changing boreal forest, Chapin, F.S., III, M. Oswood, K. Van Cleve, L.A. Viereck, and D. Verbyla (eds.). Oxford University Press, New York.

Craig, E., and F. Huettmann. 2008. Using "blackbox" algorithms such as TreeNet and Random Forests for data-mining and for finding meaningful patterns, relationships and outliers in complex ecological data: an overview, an example using golden eagle satellite data and an outlook for a promising future. in Intelligent Data Analysis: Developing New Methodologies through Pattern Discovery and Recovery, Wang, H.-f. (ed.). IGI Global, Hershey, PA, USA.

Cressie, N.A.C. 1993. Statistics for spatial data. John Wiley & Sons, New York, New York, USA.

Cushman, S., and F. Huettmann (eds.). 2010. Spatial complexity, informatics, and wildlife conservation. Springer, Tokyo, Japan. 464 p.

Cushman, S.A., and K.S. McKelvey. 2009. Data on distribution and abundance: monitoring for research and management. P. 111-129 in Spatial complexity, informatics, and wildlife conservation, Cushman, S.A., and F. Huettmann (eds.). Springer Tokyo, Japan.

Cutler, D.R., T.C. Edwards, Jr., K.H. Beard, A. Cutler, K.T. Hess, J. Gibson, and J.J. Lawler. 2007. Random forests for classification in ecology. Ecology 88(11):2783-2792.

De'ath, G., and K.E. Fabricius. 2000. Classification and regression trees: A powerful yet simple technique for ecological data analysis. Ecology 81(11):3178-3192.

Drew, C.A., Y.F. Wiersma, and F. Huettmann (eds.). 2010. *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications*. Springer Verlag, New York, New York, USA. 314 p.

ESRI. 2011. *ArcGIS Desktop: Release 10*. Environmental Systems Research Institute, Redlands, CA.

Fassnacht, K.S., W.B. Cohen, and T.A. Spies. 2006. Key issues in making and using satellite-based maps in ecology: A primer. *For. Ecol. Manage.* 222(1-3):167-181.

Ferrier, S., G. Watson, J. Pearce, and M. Drielsma. 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. *Biodivers. Conserv.* 11(12):2275-2307.

Fleming, M. 1997. A statewide vegetation map of Alaska using phenological classification of AVHRR data. P. 25-26 *in* The Second Circumpolar Arctic Vegetation Mapping Workshop Walker, D.A., and A.C. Lillie (eds.). Institute of Arctic and Alpine Research Occasional Paper Arendal, Norway, 18-24 May 1996 and the CAVM-North American Workshop, Anchorage, Alaska, US, 14-16 January 1997.

Franklin, J. 1995. Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. *Prog. Phys. Geogr.* 19(4):474-499.

Franklin, S.E. 2001. *Remote sensing for sustainable forest management*. CRC Press, Boca Raton, Fla.

Fresco, N. 2006. *Carbon sequestration in Alaska's boreal forest: Planning for resilience in a changing landscape*. PhD dissertation, University of Alaska Fairbanks, Fairbanks, Alaska.

GAO. 2005. Natural Resources. Federal Agencies Are Engaged in Various Efforts to Promote the Utilization of Woody Biomass, but Significant Obstacles to Its Use Remain. United States Government Accountability Office, Washington, DC.

Grossmann, E., J. Ohmann, J. Kagan, H. May, and M. Gregory. 2010. Mapping ecological systems with a random forest model: tradeoffs between errors and bias. USGS Gap Analysis Bulletin:16–22.

Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* 8(9):993-1009.

Harper, K., C. Boudreault, L. DeGrandpré, P. Drapeau, S. Gauthier, and Y. Bergeron. 2003. Structure, composition, and diversity of old-growth black spruce boreal forest of the Clay Belt region in Quebec and Ontario. *Environmental Reviews* 11(S1):S79-S98.

Harrell, P.A., L.L. BourgeauChavez, E.S. Kasischke, N.H.F. French, and N.L. Christensen. 1995. Sensitivity of ERS-1 and JERS-1 radar data to biomass and stand structure in Alaskan boreal forest. *Remote Sens. Environ.* 54(3):247-260.

Hothorn, T., K. Hornik, and A. Zeileis. 2006a. Party: a laboratory for recursive part(y)itioning. <http://CRAN.R-project.org/>.

Hothorn, T., K. Hornik, and A. Zeileis. 2006b. Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3):651-674.

Houghton, R.A. 2005. Aboveground forest biomass and the global carbon balance. *Glob. Change Biol.* 11(6):945-958.

Iverson, L.R., and A.M. Prasad. 2001. Potential changes in tree species richness and forest community types following climate change. *Ecosystems* 4(3):186-199.

- Jenkins, J.C., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2003. National-scale biomass estimators for United States tree species. *For. Sci.* 49(1):12-35.
- Jenkins, J.C., D.C. Chojnacky, L.S. Heath, and R.A. Birdsey. 2004. Comprehensive Database of Diameter-based Biomass Regressions for North American Tree Species. US Dept. of Agriculture, Forest Service, Northeastern Research Station.
- Jenness, J. 2006. Topographic position index (tpi\_jen.avx) extension for Arcview 3.x, v.1.3a. <http://www.jennessent.com/arcview/tpi.htm>, Jenness Enterprises [EB/OL].
- Johnson, K.D., J. Harden, A.D. McGuire, N.B. Bliss, J.G. Bockheim, M. Clark, T. Nettleton-Hollingsworth, M.T. Jorgenson, E.S. Kane, M. Mack, J. O'Donnell, C.L. Ping, E.A.G. Schuur, M.R. Turetsky, and D.W. Valentine. 2011. Soil carbon distribution in Alaska in relation to soil-forming factors. *Geoderma* 167-68(0):71-84.
- Kneeshaw, D., and S. Gauthier. 2003. Old growth in the boreal forest: A dynamic perspective at the stand and landscape level. *Environmental Reviews* 11(S1):S99-S114.
- Kohyama, T. 1993. Size-Structured Tree Populations in Gap-Dynamic Forest - the Forest Architecture Hypothesis for the Stable Coexistence of Species. *J. Ecol.* 81(1):131-143.
- Kursa, M.B., and W.R. Rudnicki. 2010. Feature Selection with the Boruta Package. *Journal of Statistical Software* 36(11):1-13.
- Lamsal, S., D.M. Rizzo, and R.K. Meentemeyer. 2012. Spatial variation and prediction of forest biomass in a heterogeneous landscape. *Journal of Forestry Research* 23(1):13-22.
- Legendre, P. 1993. Spatial Autocorrelation - Trouble or New Paradigm. *Ecology* 74(6):1659-1673.

- Li, J., A.D. Heap, A. Potter, and J.J. Daniell. 2011. Application of machine learning methods to spatial interpolation of environmental variables. *Environ Modell Softw* 26(12):1647-1659.
- Liang, J.J. 2010. Dynamics and management of Alaska boreal forest: An all-aged multi-species matrix growth model. *For. Ecol. Manage.* 260(4):491-501.
- Liang, J.J., and M. Zhou. 2010. A geospatial model of forest dynamics with controlled trend surface. *Ecol. Model.* 221(19):2339-2352.
- Liaw, A., and M. Wiener. 2002. Classification and Regression by randomForest. *R News* 2(3):18-22.
- Lloyd, A.H., and C.L. Fastie. 2002. Spatial and temporal variability in the growth and climate response of treeline trees in Alaska. *Climatic Change* 52(4):481-509.
- Loeffler, D., J. Brandt, T. Morgan, and G. Jones. 2010. Forestry Based Biomass Economic and Financial Information and Tools: An Annotated Bibliography. P. 52, Gen. Tech. Rep. RMRS-GTR-244WWW. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Major, J. 1951. A functional, factorial approach to plant ecology. *For. Ecol. Manage.* 32:392-412.
- Malone, T., J. Liang, and E.C. Packee. 2009. Cooperative Alaska Forest Inventory. Gen. Tech. Rep. PNW-GTR-785. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 42 p.
- McCarthy, J. 2001. Gap dynamics of forest trees: a review with particular attention to boreal forests. *Environmental Reviews* 9(1):1-59.

McRoberts, R.E., S. Winter, G. Chirici, E. Hauk, D.R. Pelz, W.K. Moser, and M.A. Hatfield. 2008. Large-scale spatial patterns of forest structural diversity. *Can. J. For. Res.-Rev. Can. Rech. For.* 38(3):429-438.

Miles, P.D., G.J. Brand, C.L. Alerich, L.F. Bednar, S.W. Woudenberg, J.F. Glover, and E.N. Ezell. 2001. The forest inventory and analysis database description and users manual version 1.0. Gen. Tech. Rep. NC-218. St. Paul, MN: US Department of Agriculture, Forest Service, North Central Research Station.

Miller, J., J. Franklin, and R. Aspinall. 2007. Incorporating spatial dependence in predictive vegetation models. *Ecol. Model.* 202(3-4):225-242.

Niemelä, J. 1999. Management in Relation to Disturbance in the Boreal Forest. *For. Ecol. Manage.* 115:127-134.

Ohse, B., F. Huettmann, S.M. Ickert-Bond, and G.P. Juday. 2009. Modeling the distribution of white spruce (*Picea glauca*) for Alaska with high accuracy: an open access role-model for predicting tree species in last remaining wilderness areas. *Polar Biol.* 32(12):1717-1729.

O'Neill, R.V., C.T. Hunsaker, K.B. Jones, K.H. Riitters, J.D. Wickham, P.M. Schwartz, I.A. Goodman, B.L. Jackson, and W.S. Baillargeon. 1997. Monitoring Environmental Quality at the Landscape Scale. *Bioscience* 47(8):513-519.

- Parmentier, I., R.J. Harrigan, W. Buermann, E.T.A. Mitchard, S. Saatchi, Y. Malhi, F. Bongers, W.D. Hawthorne, M.E. Leal, S.L. Lewis, L. Nusbaumer, D. Sheil, M.S.M. Sosef, K. Affum-Baffoe, A. Bakayoko, G.B. Chuyong, C. Chatelain, J.A. Comiskey, G. Dauby, J.-L. Doucet, S. Fauset, L. Gautier, J.-F. Gillet, D. Kenfack, F.N. Kouamé, E.K. Kouassi, L.A. Kouka, M.P.E. Parren, K.S.H. Peh, J.M. Reitsma, B. Senterre, B. Sonké, T.C.H. Sunderland, M.D. Swaine, M.G.P. Tchouto, D. Thomas, J.L.C.H. Van Valkenburg, and O.J. Hardy. 2011. Predicting alpha diversity of African rain forests: models based on climate and satellite-derived data do not perform better than a purely spatial model. *J. Biogeogr.* 38(6):1164-1176.
- Powell, S.L., W.B. Cohen, S.P. Healey, R.E. Kennedy, G.G. Moisen, K.B. Pierce, and J.L. Ohmann. 2010. Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: A comparison of empirical modeling approaches. *Remote Sens. Environ.* 114(5):1053-1068.
- Prasad, A.M., L.R. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems* 9(2):181-199.
- Pretzsch, H. 2005. Diversity and productivity in forests: evidence from long-term experimental plots. P. 41-64 in *Forest Diversity and Function*, Scherer-Lorenzen, M., C. Körner, and E.-D. Schulze (eds.). Springer-Verlag, Berlin Heidelberg.
- Ruefenacht, B., M.V. Finco, M.D. Nelson, R. Czaplewski, E.H. Helmer, J.A. Blackard, G.R. Holden, A.J. Lister, D. Salajanu, D. Weyermann, and K. Winterberger. 2008. Conterminous US and Alaska Forest Type Mapping Using Forest Inventory and Analysis Data. *Photogramm. Eng. Remote Sens.* 74(11):1379-1388.
- Scherer-Lorenzen, M., C. Körner, and E.-D. Schulze. 2005. *Forest diversity and function: temperate and boreal systems*. Springer Verlag, Berlin, Heidelberg. 399 p.

Schimel, D.S., J.I. House, K.A. Hibbard, P. Bousquet, P. Ciais, P. Peylin, B.H. Braswell, M.J. Apps, D. Baker, A. Bondeau, J. Canadell, G. Churkina, W. Cramer, A.S. Denning, C.B. Field, P. Friedlingstein, C. Goodale, M. Heimann, R.A. Houghton, J.M. Melillo, B. Moore, 3rd, D. Murdiyarso, I. Noble, S.W. Pacala, I.C. Prentice, M.R. Raupach, P.J. Rayner, R.J. Scholes, W.L. Steffen, and C. Wirth. 2001. Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems. *Nature* 414(6860):169-172.

Schulze, E.D., C. Wirth, D. Mollicone, and W. Ziegler. 2005. Succession after stand replacing disturbances by fire, wind throw, and insects in the dark Taiga of Central Siberia. *Oecologia* 146(1):77-88.

Stage, A.R., and C. Salas. 2007. Interactions of elevation, aspect, and slope in models of forest species composition and productivity. *For. Sci.* 53(4):486-492.

Van Cleve, K., C.T. Dyrness, L.A. Viereck, J. Fox, F.S. Chapin, and W. Oechel. 1983. Taiga ecosystems in interior Alaska. *Bioscience* 33(1):39-44.

Viereck, L., and E. Little. 2007. Alaska trees and shrubs. University of Alaska Press, Fairbanks, Alaska, USA.

Walsh, J.E., W.L. Chapman, V. Romanovsky, J.H. Christensen, and M. Stendel. 2008. Global Climate Model Performance over Alaska and Greenland. *J. Clim.* 21(23):6156-6174.

Wilmking, M., and G.P. Juday. 2005. Longitudinal variation of radial growth at Alaska's northern treeline - recent changes and possible scenarios for the 21st century. P. 282-300. Elsevier Science Bv.

Yarie, J., and S. Billings. 2002. Carbon balance of the taiga forest within Alaska: present and future. *Can. J. For. Res.-Rev. Can. Rech. For.* 32(5):757-767.



Yarie, J., E. Kane, and M.C. Mack. 2007. Aboveground biomass equations for trees of Interior Alaska. University of Alaska Fairbanks.

Yarie, J., and D. Mead. 1982. Aboveground tree biomass on productive forest land in Alaska. Res. Pap. PNW-RP-298. Portland, OR: US Department of Agriculture, Forest Service, Pacific Northwest Research Station. 16 p.

Young, B., J.J. Liang, and F.S. Chapin. 2011. Effects of species and tree size diversity on recruitment in the Alaskan boreal forest: A geospatial approach. *For. Ecol. Manage.* 262(8):1608-1617.

Table 4.1: Definition of variables used in the analysis

Variable	Description	Unit	Reference
<b>Spatial Structure</b>			
X	Northing (Alaska Albers)	10 <sup>5</sup> m	Magness et al. (2008)
Y	Easting (Alaska Albers)	10 <sup>5</sup> m	Magness et al. (2008)
<b>Climatic Variables</b>			
T_01	Mean temperature January	(°C + 100)	Ohse et al. (2009)
T_05	Mean temperature May	(°C + 100)	Ohse et al. (2009)
T_06	Mean temperature June	(°C + 100)	Ohse et al. (2009)
T_07	Mean temperature July	(°C + 100)	Ohse et al. (2009)
T_08	Mean temperature August	(°C + 100)	Ohse et al. (2009)
T_09	Mean temperature September	(°C + 100)	Ohse et al. (2009)
T_G	Mean temperature growing season (May-September)	(°C + 100)	Ohse et al. (2009)
T_D	Mean temperature difference July-January	(°C + 100)	Ohse et al. (2009)
T_A	Mean annual temperature	(°C + 100)	Ohse et al. (2009)
P_05	Precipitation sum May	mm	Ohse et al. (2009)
P_06	Precipitation sum June	mm	Ohse et al. (2009)
P_07	Precipitation sum July	mm	Ohse et al. (2009)
P_08	Precipitation sum August	mm	Ohse et al. (2009)
P_09	Precipitation sum September	mm	Ohse et al. (2009)
P_G	Precipitation sum growing season (May-September)	mm	Ohse et al. (2009)
P_W	Precipitation sum winter (October-April)	mm	Yarie (2008)
P_A	Precipitation sum annual	mm	Ohse et al. (2009)
<b>Topographic Variables</b>			
Solar	Potential maximum solar insolation	(kWH/m <sup>2</sup> )	Fu and Rich (1999)
Prod	Site productivity	unitless	Stage and Salas (2007)
Sl	Slope	percent	Stage and Salas (2007)
As	Transformed aspect	unitless	Beers et al. (1966)
El	Elevation	m	Magness et al. (2008)
SPC	Slope position classification	class	Murphy et al. (2010)
TPI	Topographic position index	class	Murphy et al. (2010)
LC	Landform classification	class	Johnson et al. (2011)
<b>Vegetation Variables</b>			
Veg	Vegetation type	class	Fleming (1997)
FT	Forest Type	class	Ruefenacht et al. (2008)
Gmax	Maximum NDVI	NDVI	Magness et al. (2008)
Gmean	Mean NDVI growing season	NDVI	Magness et al. (2008)

Table 4.2: Definition of variables used in the analysis continued

Variable	Description	Unit	Reference
Age	Stand age	Binary	Johnson et al. (2011)
H <sub>s</sub>	Tree species diversity	Shannon's	Young (chapter 3)
H <sub>d</sub>	Tree size class diversity	Shannon's	Young (chapter 3)
<b>Anthropogenic Variables</b>			
Dtc	Distance to community	km	Wurtz et al. (2006)
Dtr	Distance to roadway	km	Wurtz et al. (2006)
Dtw	Distance to navigable waterway	km	Wurtz et al. (2006)
<b>Geophysical variables</b>			
Perm	Permafrost	class	Liang (2010)
Soil	Soil type	class	Ohse et al. (2009)

**Table 4.3: Vegetation variables (Veg, from Table 4.1) used in this analysis were derived from the classification system developed by Fleming (1997).**

<b>Vegetation Value</b>	<b>Vegetation Class Names</b>
1	Low & Dwarf Shrub
2	Tall Shrub
3	Low Shrub/Lichen Tundra
4	Dwarf Shrub Tundra
5	Closed Broadleaf & Closed Mixed Forest
6	Closed Mixed Forest
7	Closed Spruce Forest
8	Spruce Woodland/Shrub
9	Open Spruce Forest/Shrub/Bog Mosaic
10	Spruce & Broadleaf Forest
11	Open & Closed Spruce Forest
12	Open Spruce & Closed Mixed Forest Mosaic
13	Tall & Low Shrub
14	Forest 1991 Fires
15	Closed Spruce & Hemlock
16	1990 Fires & Gravel Bars

**Table 4.4: Variable importance (ranking) in determining aboveground forest woody biomass (biomass; Mg/ha dry weight) using percent increase in mean standard error (%IncMSE) for ranking purposes. The variables deemed important were used in the development of the final random forest analysis (RFA).**

Variables	Biomass	
	%IncMSE	rank
H <sub>d</sub>	90.62	1
Veg	36.40	2
P_06	31.86	3
Dtc	29.80	4
El	28.79	5
Y	27.52	6
Gmax	25.68	7
Prod	24.90	8
Sl	23.87	9
P_W	23.67	10
Gmean	23.05	11
Asp	22.04	12
P_07	21.65	13
T_A	21.00	14
X	20.94	15
TPI	20.74	16
P_A	19.37	17
FT	19.32	18
P_G	18.98	19
Perm	16.71	20
Solar	16.51	21
P_09	16.43	22
T_G	16.35	23
P_08	16.17	24
T_05	15.68	25
Age	14.45	26
Soil	14.20	27
T_07	13.96	28
T_09	13.33	29
P_05	13.30	30
T_08	12.61	31
T_06	12.41	32

**Table 4.5: Spatial autocorrelation and its level of significance for aboveground forest woody biomass (biomass; Mg/ha dry weight) in the Alaskan boreal forest, and for the residuals of the random forest analysis (RFA) model used for predicting biomass from this dataset.**

	Moran's <i>I</i>	<i>P</i> -Value	Geary's <i>C</i>	<i>P</i> -Value
<b>Biomass</b>	0.3224	<0.001	0.6779	<0.001
<b>Residuals of RFA model for biomass</b>	-0.0345	0.8906	1.0202	0.7535

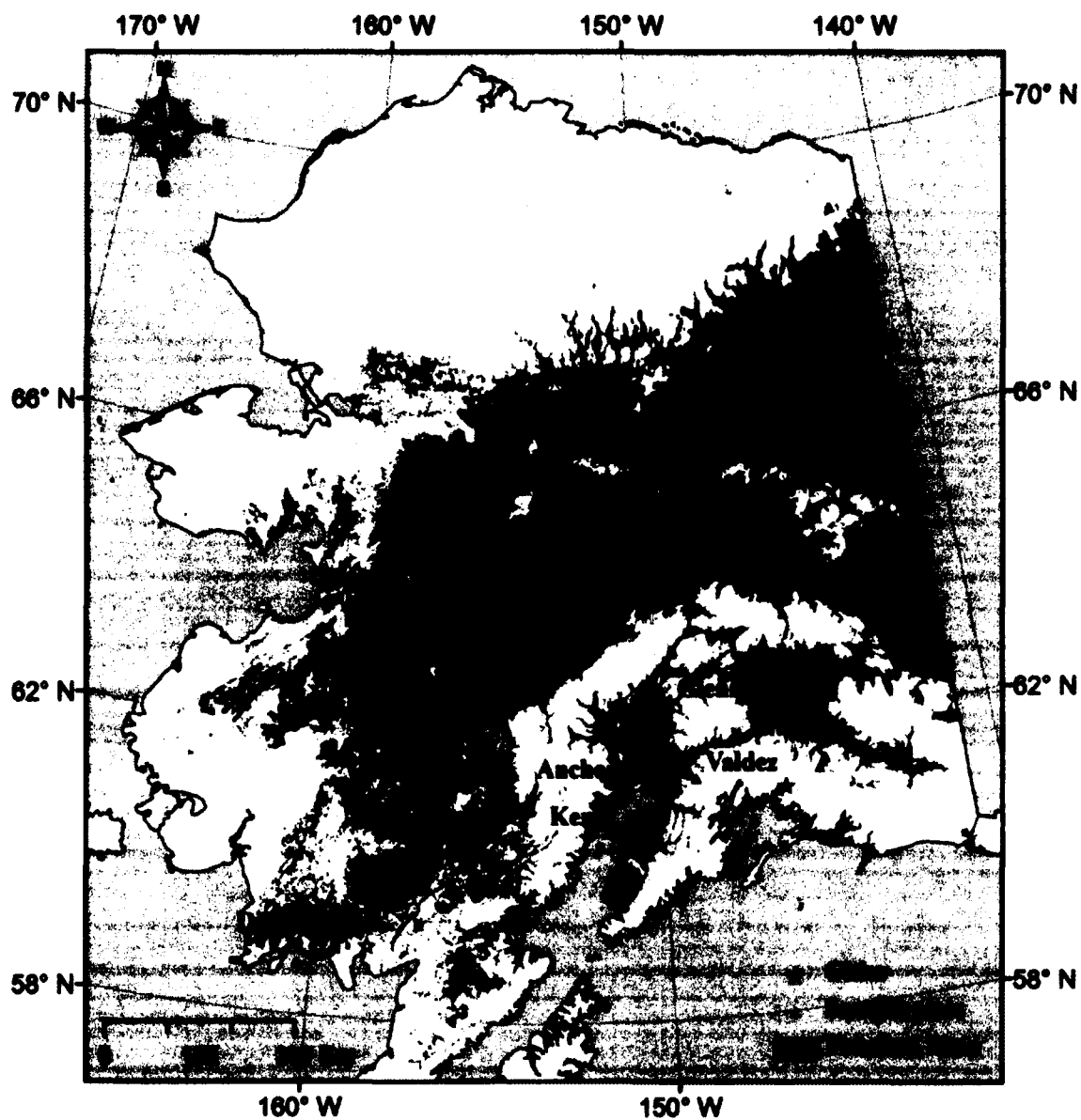


Figure 4.1: Geographic distribution of the 704 Sample Plots (in triangles) within the Alaskan boreal forest (Ruefenacht et al., 2008).

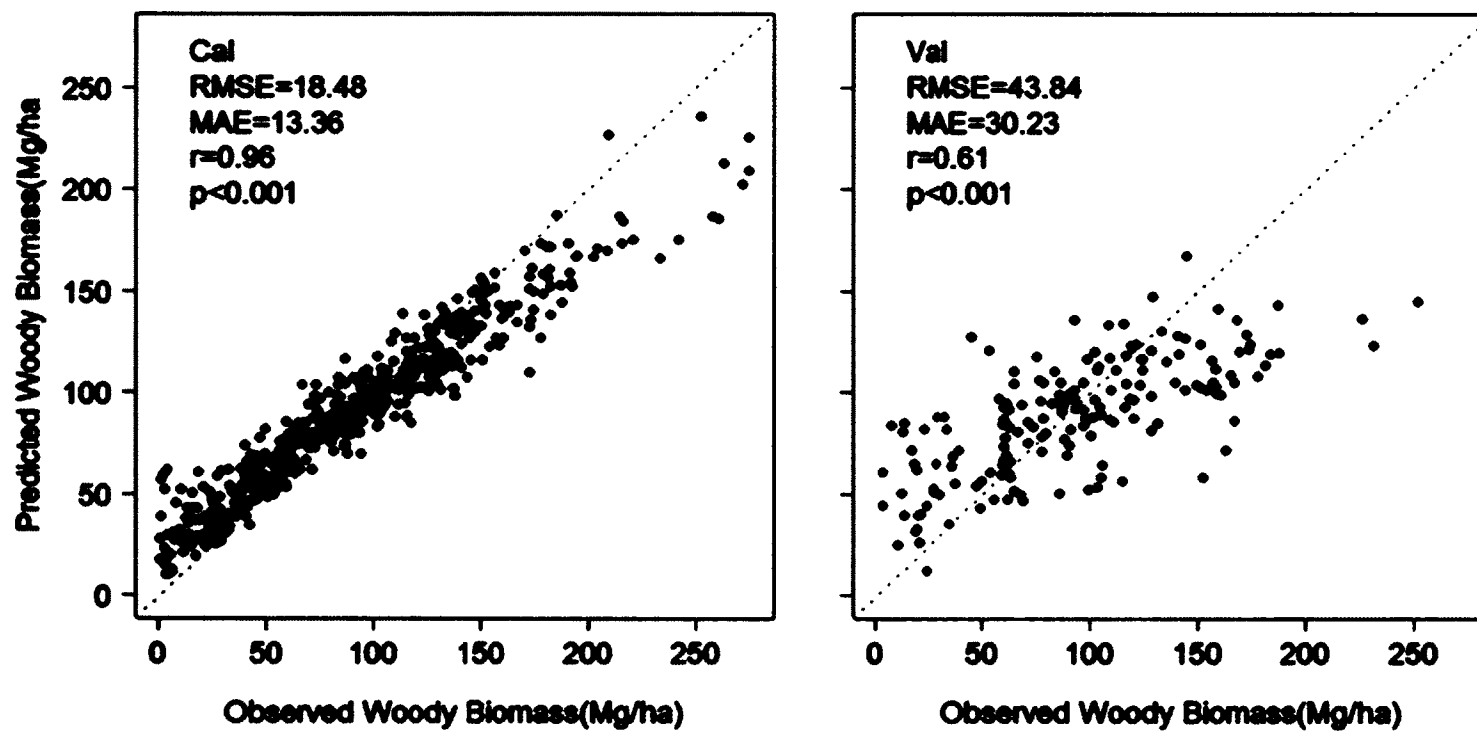


Figure 4.2: Application of the RFA model for woody biomass to the calibration data set (Cal) and the validation data set (Val).



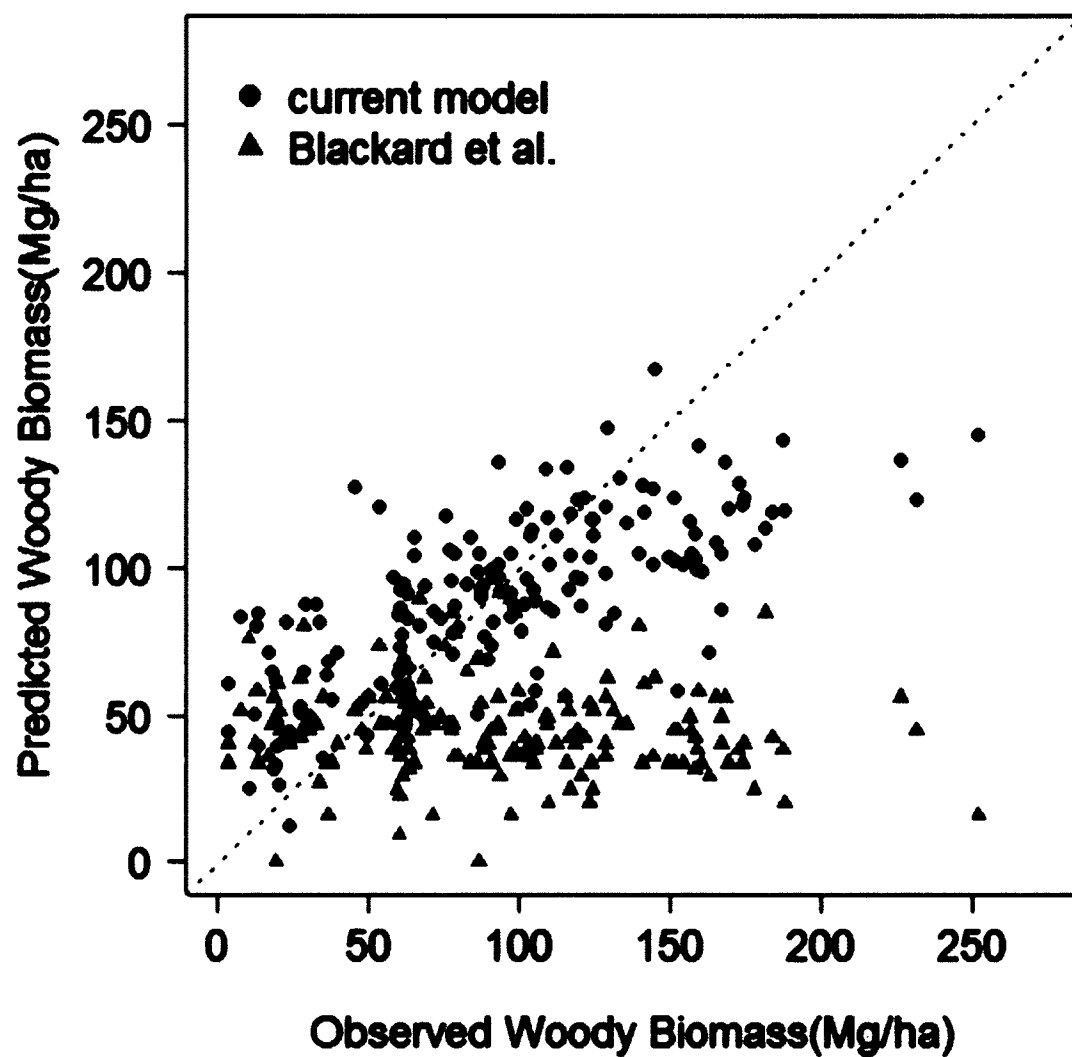


Figure 4.3: Predicted and observed above ground full tree woody biomass (Mg/ha dry weight) for the 704 Sample Plots within the Alaskan boreal forest. The current model (black dots) indicates the model presented in this paper while the other predictions are from Blackard et al. (2008, red triangles).

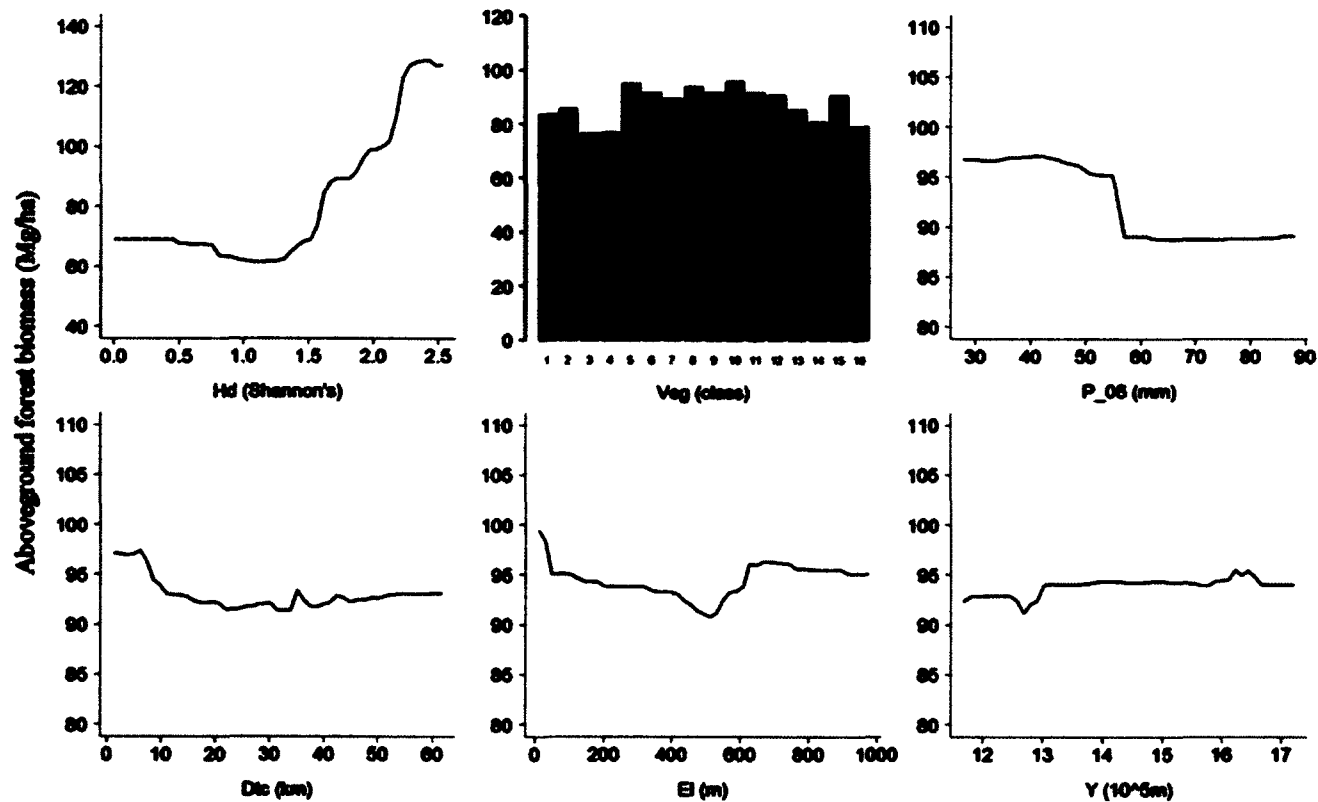


Figure 4.4: Partial dependence plots for the six most important variables from the random forest analysis (RFA; Table 4.3) predictions of aboveground forest woody biomass on tree size-class diversity (Hd), Vegetation type (Veg, refer to Table 4.2), June precipitation (P\_06), Distance to communities (Dtc), Elevation (El), and Longitude (Y). Partial dependence is the predicted value of the response based on the value of one predictor variable after averaging out the effects of the other predictor variables in the model.

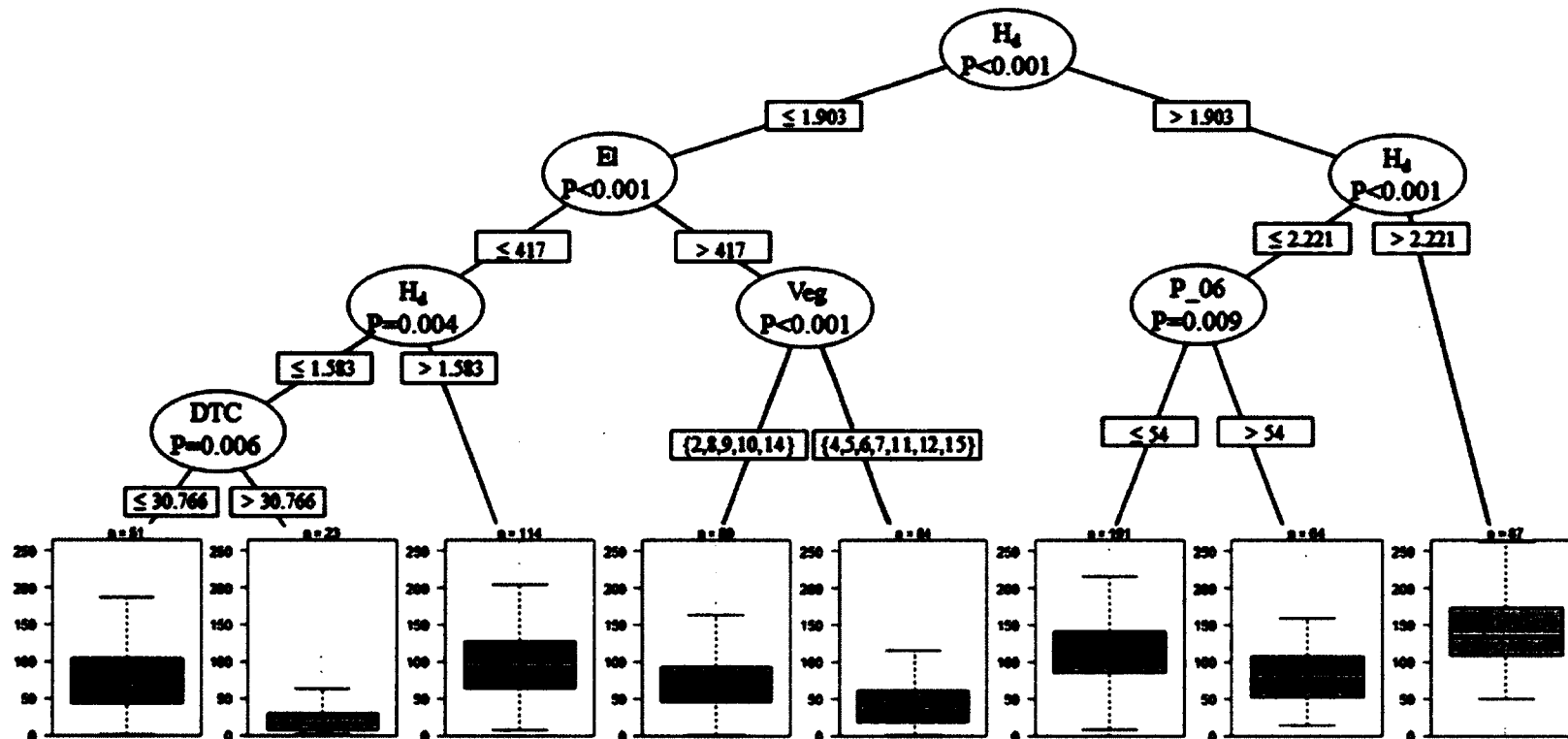


Figure 4.5: Conditional Inference tree for aboveground forest woody biomass using the six most important predictor variables as determined by the random forest analysis (see Figure 4.4, Table 4.3). The p-values at each of the nodes are from a Monte Carlo randomization test and in order for a split to occur the p-value must be  $<0.001$ . The box plots at the terminal nodes represent the distribution of the data within that branch of the tree. The shaded areas in the boxes represent the inner-quartile range while the horizontal lines represent the median values and, the whiskers extend to the most extreme data point that is no more than 1.5 times the inner-quartile range.

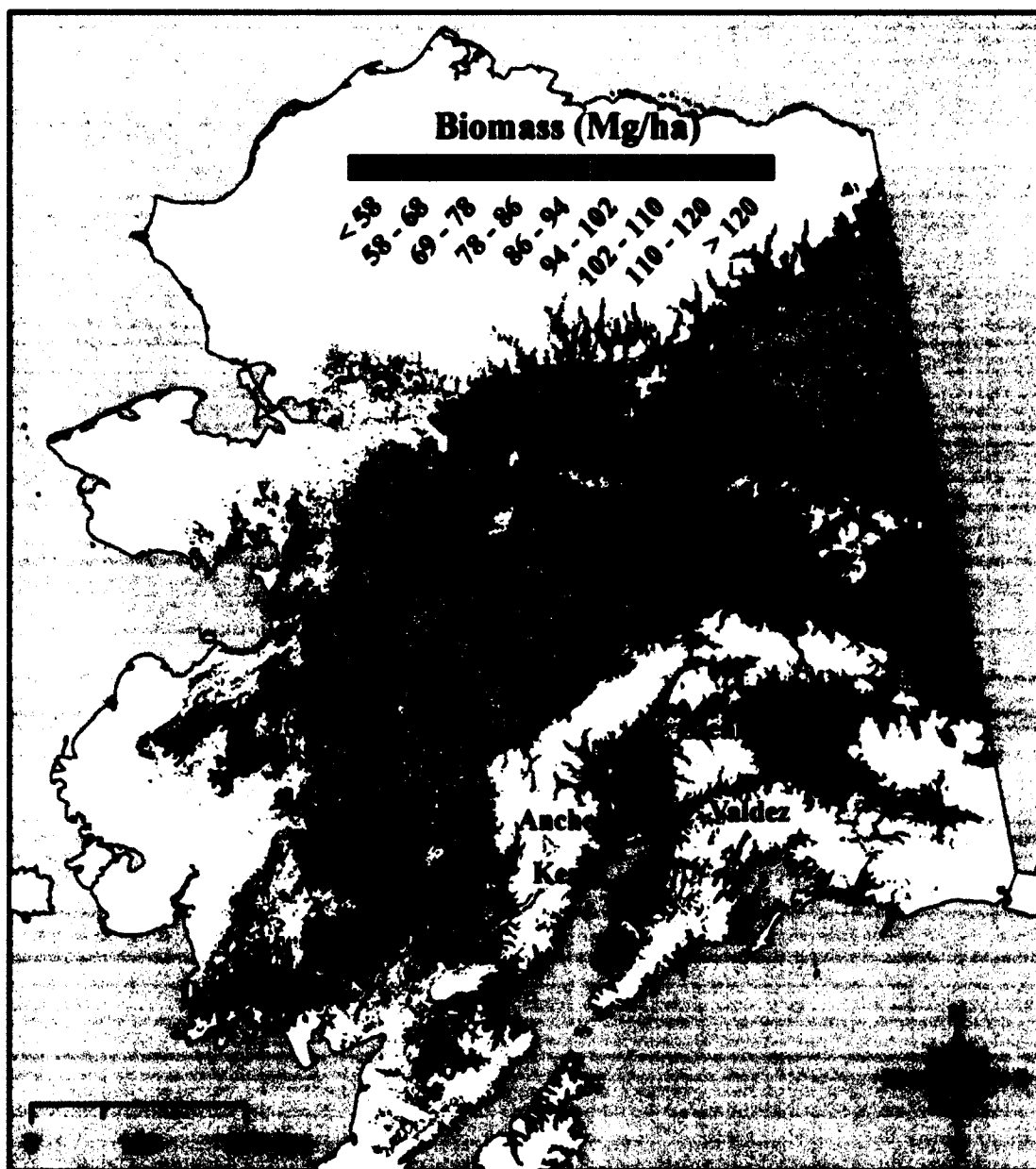


Figure 4.6: A map of the aboveground forest woody biomass (Mg/ha dry weight) derived from the random forest analysis model using the CAFI and WAIN forest inventory plot biomass values predicted across the landscape as a function of environmental predictors (Table 4.3).

### Appendix 4.1: R code used in the Analysis of Chapter 4

```

#Load The library's #
library(geoR)
library(MASS)
library(spdep)
library(RANN)
library(gstat)
library(sp)
library(maptools)
library(RColorBrewer)
library(hier.part)
library(survival)
library(VGAM)
library(rpart)
library(tree)
library(randomForest)
library(ModelMap)
library(party)
library(Boruta)

#load the data
# set seed
set.seed(71541)

#Training set data (biomasstrainingdata.csv)
mydata=read.csv(file.choose(),header=T)

#subset training data in cal and val datasets
select <- sample(1:nrow(mydata), round(nrow(mydata)*.25,0), replace =
FALSE)
dataval=mydata[select,]
datacal=mydata[-select,]

#full grid data (Predictiongriddata.csv)
datafull = read.csv(file.choose(),header=T)
datafull$Soil=as.factor(datafull$Soil)
datafull$FT=as.factor(datafull$FT)
datafull$Veg=as.factor(datafull$Veg)
datafull$Perm=as.factor(datafull$Perm)
datafull$Age=as.factor(datafull$Age)

```

```

#load functions used
#Calculate r2
r2 <- function(obs,pred) {
  SSE <- sum((obs-pred)^2)
  SST <- sum((obs-mean(obs))^2)
  return((1-SSE/SST))
}
#Calculate root mean square error (rmse)
rmse <- function(x,y) { sqrt(mean((x-y)^2,na.rm=TRUE)) }
#Calculate mean square error (mse)
mse <- function(x,y) { (mean((x-y)^2,na.rm=TRUE)) }
#calculate bias
bias <- function(x,y) { (mean((x-y),na.rm=TRUE)) }
#calculate mean absolute error (MAE)
mae <- function(x,y) { (mean(abs(x-y),na.rm=TRUE)) }

#convert coordinates data to matrix to be used for neighborhood structure
xy=as.matrix(as.data.frame(cbind(mydata$X*100000,mydata$Y*100000)))

#identify neighbors
k1a=knearneigh(xy,k=10,longlat=NULL,RANN=TRUE)
all.linked1=max(unlist(nbdists(k1a,xy,longlat=NULL)))
dist=30500
nb1=dnearneigh(xy, d1=0, d2=dist, row.names=NULL,longlat=NULL)
summary(nb1,xy,longlat=NULL)

#plot neighbor data
plot(xy)
plot(nb1,xy, add=TRUE,col="red", lty=2, cex=1)

#convert neighbor matrix to weight matrix
dlist1=nbdists(nb1,xy)
idlist1=lapply(dlist1,function(x)1/x)
nbw1=nb2listw(nb1,glist=idlist1,style="W", zero.policy=TRUE)
summary(nbw1)

#tests for autocorrelation for Hd
moranHd=moran.mc(mydata$Woody.Biomass..mg.ha.,nbw1,zero.policy=TRUE,
nsim=9999)
moranHd

gearyHd=geary.mc(mydata$Woody.Biomass..mg.ha.,nbw1,zero.policy=TRUE,
nsim=9999)
gearyHd

```

#tests for autocorrelation in RF final model for Biomass

```
lmnetmoran=moran.mc(resid_Woody.Biomass..mg.ha._final,nbw1,zero.policy=TRUE,
nsim=9999)
lmnetmoran
```

```
lmnetgeary=geary.mc(resid_Woody.Biomass..mg.ha._final,nbw1,zero.policy=TRUE,
nsim=9999)
lmnetgeary
```

#Using Random forest Biomass

```
Biomass_step <- Boruta(Woody.Biomass..mg.ha. ~ X + Y + Age + Solar + Prod
+ maxAV + meanAV + Hd + Hs + T_01 + T_05 + T_06 + T_07 + T_08 + T_09 +
T_G + T_D + T_A + P_05 + P_06 + P_07 + P_08 + P_09 + P_G + P_A + P_W +
El + Asp + Sl + Soil + Dtc + Dtw + Dtr + Veg + Perm + FT + LC + TPI +
SPC,data=mydata,maxRuns = 500,doTrace=2)
```

```
print(Biomass_step)
```

```
plot(Biomass_step,whichRand=c(FALSE,FALSE,FALSE),cex.lab=1.5,cex.axis=
1.5)
```

```
getConfirmedFormula(Biomass_step)
```

#tuning RF model for optimal mtry parameter

```
response= as.vector(mydata$Woody.Biomass..mg.ha.)
```

```
predictors = as.matrix(cbind(mydata$X,mydata$Y,mydata$Age,mydata$Prod,
mydata$maxAV,mydata$meanAV,mydata$Hd,mydata$T_05,mydata$T_06,mydata$T_07,mydata$T_08,mydata$T_G,mydata$T_A,mydata$P_05,mydata$P_06,mydata$P_07,mydata$P_08,mydata$P_09,mydata$P_G,mydata$P_A,mydata$P_W,mydata$El,mydata$Asp,mydata$Sl,mydata$Soil,mydata$Dtc,mydata$Veg,mydata$Perm,mydata$FT,mydata$TPI,mydata$Solar,mydata$T_09))
```

```
tuneRF(predictors,response, stepfactor=0.1,ntreeTry=100)
```

#run final RF model

```
Biomass <- randomForest(Woody.Biomass..mg.ha. ~ X + Y + Age + Prod +
maxAV + meanAV + Hd + T_05 + T_06 + T_07 + T_08 + T_G + T_A + P_05 +
P_06 + P_07 + P_08 + P_09 + P_G + P_A + P_W + El + Asp + Sl + Soil +
Dtc + Veg + Perm + FT + TPI + Solar + T_09,
data=mydata,ntree=2500,importance=TRUE)
```

```

round(importance(Biomass,type=1),2)
predict_Biomass=predict(Biomass,mydata)
rmse(predict_Biomass,mydata$Woody.Biomass..mg.ha.)
bias(predict_Biomass,mydata$Woody.Biomass..mg.ha.)
cor.test(predict_Biomass,mydata$Woody.Biomass..mg.ha.,method="pearson")
mae(predict_Biomass,mydata$Woody.Biomass..mg.ha.)
mean(mydata$Woody.Biomass..mg.ha.)
mean(Biomass$mse)
mean(Biomass$rsq)
resid_Woody.Biomass..mg.ha._final=predict_Biomass -
mydata$Woody.Biomass..mg.ha.
plot(resid_Woody.Biomass..mg.ha._final~predict_Biomass)
plot(Biomass)
rmse(mydata$FIABio,mydata$Woody.Biomass..mg.ha.)
bias(mydata$FIABio,mydata$Woody.Biomass..mg.ha.)
cor.test(mydata$FIABio,mydata$Woody.Biomass..mg.ha.,method="pearson")
mae(mydata$FIABio,mydata$Woody.Biomass..mg.ha.)
mean(mydata$FIABio)

```

#### # conditional inference tree

```

tree_biomass <- ctree(Woody.Biomass..mg.ha. ~ X + Y + Age + Prod + maxAV
+ meanAV + Hd + T_05 + T_06 + T_07 + T_08 + T_G + T_A + P_05 + P_06 +
P_07 + P_08 + P_09 + P_G + P_A + P_W + El + Asp + Sl + Soil + Dtc + Veg +
Perm + FT + TPI + Solar + T_09,
data=mydata),controls=ctree_control(mincriterion=0.99))

```

```

tree_biomass <- ctree(Woody.Biomass..mg.ha. ~ Hd + Veg + P_06 + Dtc + El +
Y, data=mydata,controls=ctree_control(mincriterion=0.99))

```

```

print(tree_biomass)

```

```

plot(tree_biomass,cex=1.5,inner_panel = node_inner(tree_biomass,id =
FALSE),terminal_panel = node_boxplot(tree_biomass,
width = 0.75,ylines = 2,yscale=c(0,265),cex=0,id= FALSE))

```

#### #analyze final model using calibration and validation data sets

```

Biomass.cal <- randomForest(Woody.Biomass..mg.ha. ~ X + Y + Age + Prod +
maxAV + meanAV + Hd + T_05 + T_06 + T_07 + T_08 + T_G + T_A + P_05 +
P_06 + P_07 + P_08 + P_09 + P_G + P_A + P_W + El + Asp + Sl + Soil + Dtc +
Veg + Perm + FT + TPI + Solar + T_09, data=datacal,ntree=2500,
importance=TRUE)

```

```

round(importance(Biomass.cal,type=1),2)

```



```

predict_Biomass.cal=predict(Biomass.cal,datacal)
rmse(predict_Biomass.cal,datacal$Woody.Biomass..mg.ha.)
bias(predict_Biomass.cal,datacal$Woody.Biomass..mg.ha.)
cor.test(predict_Biomass.cal,datacal$Woody.Biomass..mg.ha.,method="pearson")
mae(predict_Biomass.cal,datacal$Woody.Biomass..mg.ha.)
mean(Biomass.cal$mse)
mean(Biomass.cal$rsq)
resid_Woody.Biomass..mg.ha._final=predict_Biomass.cal -
datacal$Woody.Biomass..mg.ha.
plot(resid_Woody.Biomass..mg.ha._final~predict_Biomass.cal)
predict_Biomass_val=predict(Biomass.cal,dataval)
rmse(predict_Biomass_val,dataval$Woody.Biomass..mg.ha.)
bias(predict_Biomass_val,dataval$Woody.Biomass..mg.ha.)
cor.test(predict_Biomass_val,dataval$Woody.Biomass..mg.ha.,method="pearson"
)
mae(predict_Biomass_val,dataval$Woody.Biomass..mg.ha.)

```

#plot calibration and validation results

```
par(mar = c(5.1,6.1,4.1,2.1)+0.1, mgp=c(3.5,1,0))
```

```

plot(predict_Biomass.cal~datacal$Woody.Biomass..mg.ha.,xlab="Observed
Woody
Biomass(Mg/ha)",ylab="",xlim=c(0,275),ylim=c(0,275),cex.axis=1.5,cex.lab=1.5,
las=1,pch=19) mtext("Predicted Woody Biomass(Mg/ha)",cex=1.5,side=2,line=4)

```

```

myline.fit.cal <- lm(predict_Biomass.cal~datacal$Woody.Biomass..mg.ha.)
summary(myline.fit.cal)
abline(myline.fit.cal,col="RED",cex=1.5,lwd=2)
abline(a=0,b=1,lty="dotted",col="gray70",cex=1.5,lwd=2)
legend("topleft",c("Cal","RMSE=18.48","MAE=13.36","r=0.96","p<0.001"),
col=c("black"), bty = "n",cex=1.5)

```

```

plot(predict_Biomass_val~dataval$Woody.Biomass..mg.ha.,xlab="Observed
Woody
Biomass(Mg/ha)",ylab="",xlim=c(0,275),ylim=c(0,275),cex.axis=1.5,cex.lab=1.5,
las=1,pch=19) mtext("Predicted Woody Biomass(Mg/ha)",cex=1.5,side=2,line=4)

```

```

myline.fit.val <- lm(predict_Biomass_val~dataval$Woody.Biomass..mg.ha.)
summary(myline.fit.val)
abline(myline.fit.val,col="RED",cex=1.5,lwd=2)
abline(a=0,b=1,lty="dotted",col="gray70",cex=1.5,lwd=2)
legend("topleft",c("Val","RMSE=43.84","MAE=31.39","r=0.55","p<0.001"),
col=c("black"), bty = "n",cex=1.5)

```

```
# comparison b/w RFA model and Blackard et al.
```

```
par(mar = c(5.1,6.1,4.1,2.1)+0.1, mgp=c(3.5,1,0))
```

```
plot(predict_Biomass_val~dataval$Woody.Biomass..mg.ha.,xlab="Observed  
Woody
```

```
Biomass(Mg/ha)",ylab="",xlim=c(0,275),ylim=c(0,275),cex.axis=1.5,cex.lab=1.5,  
las=1, pch=19) mtext("Predicted Woody  
Biomass(Mg/ha)",cex=1.5,side=2,line=4)
```

```
points(dataval$FIABio~dataval$Woody.Biomass..mg.ha.,col="RED",pch=17)
```

```
myline.fit <- lm(predict_Biomass_val~dataval$Woody.Biomass..mg.ha.)
```

```
summary(myline.fit)
```

```
abline(myline.fit,col="black",cex=1.5,lwd=2)
```

```
myline.fit_FIA <- lm(dataval$FIABio~dataval$Woody.Biomass..mg.ha.)
```

```
summary(myline.fit_FIA)
```

```
abline(myline.fit_FIA,col="RED",cex=1.5,lwd=2)
```

```
abline(a=0,b=1,lty="dotted",col="gray70",cex=1.5,lwd=2)
```

```
legend("topleft", c("current model", "Blackard et al."), pch=c(19,17),
```

```
lty=c(1,1),col=c("black","red"), bty = "n",inset=0.03,cex=1.5)
```

```
#partial dependence plots
```

```
par(mar = c(5,5,4,2)+ 0.1, mgp=c(3.75,1,0))
```

```
imp <- importance(Biomass) # get the importance measures
```

```
impvar <- rownames(imp)[order(imp[, 1], decreasing=TRUE)]
```

```
# get the sorted names
```

```
op <- par(mfrow=c(2, 3))
```

```
for (i in seq_along(impvar)) {
```

```
  partialPlot(Biomass, mydata, impvar[i], xlab=impvar[i],
```

```
  main=paste("Partial Dependence on", impvar[i]), ylim=c(0,120))
```

```
}
```

```
par(op)
```

```
partialPlot(Biomass,mydata,Hd,main="",ylim=c(40,140),
```

```
ylab="",xlab="Hd",cex.lab=1.75,cex.axis=1.75,las=1,lwd=2,bty="l",rug=FALSE)
```

```
partialPlot(Biomass,mydata,Veg,main="",ylim=c(0,120),ylab="",xlab="Veg",cex.
```

```
lab=1.75,cex.axis=1.75,las=1,lwd=2,bty="l",rug=FALSE)
```

```
partialPlot(Biomass,mydata,P_06,main="",ylim=c(80,110),ylab="",xlab="P_06",
```

```
cex.lab=1.75,cex.axis=1.75,las=1,lwd=2,bty="l",rug=FALSE)
```

```
partialPlot(Biomass,mydata,Dtc,main="",ylim=c(80,110),ylab="",xlab="Dtc",cex
```

```
.lab=1.75,cex.axis=1.75,las=1,lwd=2,bty="l",rug=FALSE)
```

```
partialPlot(Biomass,mydata,El,main="",ylim=c(80,110),
ylab="",xlab="El",cex.lab=1.75,cex.axis=1.75,las=1,lwd=2,bty="l",rug=FALSE)
```

```
partialPlot(Biomass,mydata,Y,main="",ylim=c(80,110),ylab="",xlab="Y",cex.lab
=1.75,cex.axis=1.75,las=1,lwd=2,bty="l",rug=FALSE)
```

```
#predict values for all locations for Biomass!
```

```
AKpredict_Biomass <- predict(Biomass,datafull)
```

```
Biomass_spat =
```

```
cbind(datafull$X*100000,datafull$Y*100000,AKpredict_Biomass)
```

```
Biomass_spat = as.data.frame(Biomass_spat)
```

```
names(Biomass_spat)=c("X","Y","Biomass")
```

```
write.csv(Biomass_spat,"Biomass_spat.csv",row.names=FALSE)
```

## Appendix 4.2: Aboveground forest woody biomass (Mg/ha dry weight) within the boreal forest of Alaska Raster Dataset Metadata Standard Version FGDC-STD-001-1998

### Identification

#### CITATION

##### CITATION INFORMATION

PUBLICATION DATE 2011-04-06

PUBLICATION TIME 000000

##### TITLE

Aboveground forest woody biomass (Mg/ha dry weight) within the boreal forest of Alaska

GEOSPATIAL DATA PRESENTATION FORM raster digital data

#### DESCRIPTION

##### ABSTRACT

This raster layer is from a study that modeled and mapped forest biomass within the boreal forest of Alaska. The modeling effort employed the Boruta Algorithm, Random Forest analysis, and Regression Tree analysis (three different machine learning techniques) to predict forest biomass for the entire region using a combination of forest inventory data and a suite of 32 predictors from public open-access data archives that included spectral reflectance, climatic, soil, distance from various features, and topographic variables.

The region of interest for this study was the Alaskan boreal forest which extends from the Bering Sea on the west to the Canadian border in the east and is bounded in the north by the Brooks Range and in the south by the Alaska Range and coastal mountains.

The abstract for the paper in which this data was originally published is below for reference:

Title: Mapping and predicting aboveground biomass of trees using forest inventory data and public environmental variables within the Alaskan boreal forest

Brian Young<sup>1</sup>,

John Yarie<sup>1</sup>,

David Verbyla<sup>1</sup>,

Falk Huettmann<sup>2</sup>,

F. Stuart Chapin III<sup>3</sup>

<sup>1</sup>Department of Forest Sciences, University of Alaska Fairbanks, Fairbanks, AK 99775-7200, USA

<sup>2</sup>Institute of Arctic Biology, Biology & Wildlife Department – EWHALE lab – University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

<sup>3</sup>Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99775-7000, USA

**Abstract:** A method for mapping the predicted forest biomass was developed and tested on a study region in the boreal forest of interior Alaska. In order to understand above ground biomass values within this forest, we employed the Boruta Algorithm, Random Forest analysis, and Regression Tree analysis (three different machine learning techniques), as well as tests for spatial autocorrelation, to predict above ground woody biomass for the entire region using a combination of forest inventory data and a suite of 32 predictors from public open-access data archives that included spectral reflectance, climatic, soil, distance from various features, and topographic variables. We also developed first time high resolution prediction maps at a 1km<sup>2</sup> pixel size for aboveground woody biomass. The method employed here yielded good accuracy for the huge Alaskan landscape despite a rather limited training set. The results indicate that the geographic patterns of biomass are strongly influenced by the tree size class diversity of a given stand.

#### **PURPOSE**

Modeling of forest biomass across the boreal forest of Alaska

#### **STATUS**

MAINTENANCE AND UPDATE FREQUENCY None planned

#### **SPATIAL DOMAIN**

##### **BOUNDING COORDINATES**

WEST BOUNDING COORDINATE -162.94478

EAST BOUNDING COORDINATE -137.944993

NORTH BOUNDING COORDINATE 68.98776

SOUTH BOUNDING COORDINATE 59.094566

#### **KEYWORDS**

##### **THEME**

THEME KEYWORD THESAURUS None

THEME KEYWORD Predictive mapping; Forest biomass; Machine learning (random forest); Alaska

PLACE  
 PLACE KEYWORD THESAURUS None  
 PLACE KEYWORD Alaska; Boreal forest

ACCESS CONSTRAINTS  
 None

USE CONSTRAINTS  
 None

NATIVE DATA SET ENVIRONMENT  
 Microsoft Windows 7 Version 6.1 (Build 7601) Service Pack 1; ESRI ArcGIS  
 10.0.3.3600

### **Spatial Data Organization**

DIRECT SPATIAL REFERENCE METHOD Raster

RASTER OBJECT INFORMATION  
 RASTER OBJECT TYPE Pixel  
 ROW COUNT 1363  
 COLUMN COUNT 1215

### **Spatial Reference**

HORIZONTAL COORDINATE SYSTEM DEFINITION  
 PLANAR  
 MAP PROJECTION  
 MAP PROJECTION NAME NAD 1983 Alaska Albers  
 ALBERS CONICAL EQUAL AREA  
 STANDARD PARALLEL 55.0  
 STANDARD PARALLEL 65.0  
 LONGITUDE OF CENTRAL MERIDIAN  
 LATITUDE OF PROJECTION ORIGIN  
 FALSE EASTING 0.0  
 FALSE NORTHING 0.0

PLANAR COORDINATE INFORMATION  
 PLANAR COORDINATE ENCODING METHOD coordinate pair  
 COORDINATE REPRESENTATION  
 ABSCISSA RESOLUTION 0.0000000030536018158500163  
 ORDINATE RESOLUTION 0.0000000030536018158500163  
 PLANAR DISTANCE UNITS Meter

GEODETIC MODEL  
HORIZONTAL DATUM NAME D North American 1983  
ELLIPSOID NAME GRS 1980  
SEMI-MAJOR AXIS 6378137.0  
DENOMINATOR OF FLATTENING RATIO 298.257222101

### **Metadata Reference**

METADATA DATE 2012-08-14  
METADATA CONTACT  
CONTACT INFORMATION  
CONTACT ORGANIZATION PRIMARY  
CONTACT ORGANIZATION University of Alaska Fairbanks  
CONTACT PERSON Brian Young  
CONTACT ELECTRONIC MAIL ADDRESS bdyoung@alaska.edu  
  
METADATA STANDARD NAME FGDC Content Standard for Digital  
Geospatial Metadata  
METADATA STANDARD VERSION FGDC-STD-001-1998  
METADATA TIME CONVENTION local time

## CHAPTER 5: CONCLUSION

Because a gridded USDA Forest Inventory and Analysis (FIA) program does not exist for interior Alaska, I made use of the permanent sample plot (PSP's) data within the CAFI and WAIN datasets (Malone et al. 2009; Rees, personal communication). Although these data represent the largest compiled collection of information on forest dynamics in Boreal Alaska (Malone et al. 2009; Rees, personal communication) they are not uniformly dispersed across the study region. Therefore the analyses presented in this dissertation incorporated the uses of multiple spatial prediction techniques combined with accounting for spatial autocorrelation through both classical statistical techniques and machine learning algorithms (Bivand et al. 2008; Sokal and Oden 1978; Wagner and Fortin 2005). Across the heterogeneous landscapes of boreal Alaska, a dense but localized sampling effort made it possible to detect localized patterns. The predictions for more distant locations were based strictly on correlation, which typically proves to be a very powerful approach and when applied on a landscape-scale and where autocorrelation is addressed. The methods proposed in these studies should act as models and templates to assist forest managers and researchers interested in investigating forest dynamics. The data are available for them and model predictions can be assessed and further improved.

In addition to the data contained in the CAFI and WAIN datasets, this project could not have occurred without the open access data from Bonanza Creek LTER ([www.lter.uaf.edu/data](http://www.lter.uaf.edu/data)), Scenarios Network for Alaska Planning (SNAP; [www.snap.uaf.edu](http://www.snap.uaf.edu)), Alaska Geospatial Data Clearinghouse (AGDC; [agdc.usgs.gov/agdc.html](http://agdc.usgs.gov/agdc.html)), FSGeodata Clearinghouse ([fsgeodata.fs.fed.us](http://fsgeodata.fs.fed.us)), Geographic Information Network of Alaska database (GINA; [gina.alaska.edu/data](http://gina.alaska.edu/data)), and the Alaska Interagency Coordination Center ([afsmaps.blm.gov](http://afsmaps.blm.gov)). While other open access datasets exist for this study region i.e. Alaska Gap Analysis Project (AK-GAP; [aknhp.uaa.alaska.edu](http://aknhp.uaa.alaska.edu)), Alaska Mapped Statewide Digital Mapping Initiative (SDMI; [www.alaskamapped.org/sdmi](http://www.alaskamapped.org/sdmi)), and the USGS Landfire Data Distribution Site ([landfire.cr.usgs.gov](http://landfire.cr.usgs.gov)), I did not use the data contained within them because they did not yet meet my specific needs (summarized in Table 5.1). Other spatial data for this region



do exist but, it is either expensive to obtain, highly fragmented, localized, or difficult to make sense of due to a lack of clear metadata.

A well-designed spatial program that emphasizes clear standards for integrated management with open access data and their easily to understand metadata would greatly increase the utility and cost-effectiveness of any monitoring program so key knowledge gaps can begin to be closed. The ability to clearly synthesize data and decipher code has become a vital skill and a best professional practice for the modern scientist given the voluminous amounts of available data. The combination of open source through avenues such as The R Project (Team 2010) and open-access data will aid in the development of new and novel investigations and transparent science-based management (Huettmann, 2005).

Adaptive management is designed as a series of experiments to test and evaluate management alternatives and it requires a focus on learning and flexible policy (Gregory et al. 2006; Holling 1978; Walters 1986; Walters and Holling 1990). Although an adaptive management framework allows for the testing of methods for forest management, it does not allow for mutually accepted goal setting. This creation and meeting of management goals is crucial though for the perceived success of any management initiative and therefore requires collaborative adaptive management which incorporates public participation (Sulak and Huntsinger 2012). For example, within the State of Alaska, prior to implementing any forestry activity on State lands, the Division of Forestry is required to have these activities vetted through agency and public review. Much of such processes are either tradition or based on policy that does not accommodate adaptive management. Often, as a component of the public review process, these activities are discussed and brought before the Board of Forestry, a State-wide entity composed of appointed officials, and a more regional-scale community advisory council. This public dialogue and review process typically incorporates the use of maps as a cornerstone in the participatory process.

For effective adaptive forest management the strategy needs to be designed in such a way as to pursue the best possible expected overall outcome in terms of specific

performance measures. The problem lies, however, in which performance measure a given agency chooses to pursue. For example, one objective that is common in interior Alaska is to minimize the risk associated with wildfire surrounding communities (Chapin et al. 2008; Ogden and Innes 2007) through the use of shear-blading of fire breaks. An agency may also choose to have a multi-criteria objective (Schwenk et al. 2012; Zhou et al. 2008), such as salvage logging within a burn area, to reduce the spread of an insect outbreak while capturing the maximum value of the associated timber in addition to enhancing wildlife habitat. To meet any one of these objectives, or multiple objectives, clear representative descriptions of forest ecosystems and socioeconomic conditions are crucial (Chapin et al. 2006; Chapin et al. 2008). Therefore, specific models and the creation of publicly available maps describing and showing stand-to-landscape-level dynamics are required so that informed decision making can occur.

Predictive mapping is a powerful tool for landscape-level planning and analysis (Drew et al. 2010; Franklin 1995). However, predictive models and their performance can be sensitive to sample size and information contained within the samples. Consequently, landscape and regional scale predictive models often strike a balance between sample size and prediction performance. Large sample sizes can be cost-prohibitive, particularly in remote locations such as Alaska. This question of required sampling effort versus predicted accuracy and low variance is a key topic for virtually all natural resource inventories. However, the use of spatial interpolation and machine learning can drastically enhance our capacity to overcome these limitations.

The predictive mapping results presented in this dissertation can be best used for broad landscape-level planning. I believe that these models could be improved by adding more ground-sampled data, open access data sources, and a stronger collaboration with forest practitioners to assess whether these models meet their particular needs, interests, or applications. Improved understanding of the current and potential future landscape-level vegetation patterns will assist land management agencies in their decision-making processes in regards to sustainable forestry activities (Ogden and Innes 2009).

Spatially continuous data play a significant role in planning, risk assessment and decision making in environmental management and conservation. They are, however, usually not easily available and often difficult and expensive to acquire, especially for remote regions such as Alaska. Data on environmental properties are usually collected by point sampling, and this is problematic for environmental managers because they typically require spatially continuous data over a region of interest to make informed decisions. Therefore, the application of different statistical approaches is required. In this dissertation, I incorporated the use of both classical statistical approaches and machine learning (ML) to understand the diversity and its impacts on recruitment and production. The application of ML in ecology has increased in recent years (Olden et al. 2008) but many in the field are still unfamiliar with the technique and thus skeptical and on occasion even discriminatory of the approach and the results (Huettmann pers. comm.). My results show that the combination of classical statistics and machine learning algorithms can complement many studies when predictions of ecological phenomena are the overarching goals.

To achieve sustainable forest management we need “to emulate nature in our interventions in such a way as to minimize potential impacts and to conserve or enhance biodiversity” (Messier and Kneeshaw, 1999). To adhere to this concept of sustainable forest management, the State of Alaska needs to incorporate the key provision of biodiversity enhancement or at the very least maintenance into the State’s Forest Resources and Practices Act (FRPA, AS 41.17). This gesture would bring Alaska to the forefront of science (as needed for a science-based management) and into agreement with the vast majority of the other States of the United States and with numerous international laws (i.e., Convention on the Conservation of Biodiversity, CITES, The Bonn Convention). Such a move would further be in alignment with Article VIII of the State of Alaska’s constitutional mandate on sustainable yield in that diminishment of biodiversity not only impacts the provisioning ecosystem goods and services but potentially cultural, intellectual, aesthetic and spiritual values that are important to society (Chapin et al. 2000).

With a growing interest for forest products, mainly in the form of biomass, at home and a potential increase in future demand from Asia (Bormann et al. 2007), the forest industry in Interior Alaska is on the cusp of change. The potential impact of forest management on biodiversity is likely to depend on how harvesting activities differ from the disturbances to which the various species adapted over evolutionary time. While the future scale of timber harvesting may represent novel disturbances in this forest, the effects may negatively impact ecosystem services provided by the forest if they do not effectively mimic natural disturbance-recovery processes (Johnson et al. 1999). Forest management activities that help to restore forests altered by past management actions or natural disturbances may also have positive effects on native biodiversity. Forest management needs to incorporate an adaptive management framework so that sustainable forestry within the boreal forest of interior Alaska can continue to be the reality. This can be achieved by forest managers and researchers continually applying a wide array of silvicultural approaches so that the lessons learned can be applied to provide a full range of ecosystem goods and services (Schwenk et al. 2012).

## 5.1 REFERENCES

Bivand, R.S., E.J. Pebesma, and V. Gómez-Rubio. 2008. *Applied spatial data analysis with R*. Springer Verlag, New York, New York, USA.

Bormann, B.T., R.W. Haynes, and J.R. Martin. 2007. Adaptive Management of Forest Ecosystems: Did Some Rubber Hit the Road? *Bioscience* 57(2):186-191.

Chapin, F.S., 3rd, E.S. Zavaleta, V.T. Eviner, R.L. Naylor, P.M. Vitousek, H.L. Reynolds, D.U. Hooper, S. Lavorel, O.E. Sala, S.E. Hobbie, M.C. Mack, and S. Diaz. 2000. Consequences of changing biodiversity. *Nature* 405(6783):234-242.

Chapin, F.S., 3rd, A.L. Lovecraft, E.S. Zavaleta, J. Nelson, M.D. Robards, G.P. Kofinas, S.F. Trainor, G.D. Peterson, H.P. Huntington, and R.L. Naylor. 2006. Policy strategies to address sustainability of Alaskan boreal forests in response to a directionally changing climate. *Proc Natl Acad Sci U S A* 103(45):16637-16643.

Chapin, F.S., 3rd, S.F. Trainor, O. Huntington, A.L. Lovecraft, E. Zavaleta, D.C. Natcher, A.D. McGuire, J.L. Nelson, L. Ray, M. Calef, N. Fresco, H. Huntington, T.S. Rupp, L. Dewilde, and R.L. Naylor. 2008. Increasing wildfire in Alaska's boreal forest: Pathways to potential solutions of a wicked problem. *Bioscience* 58(6):531-540.

Drew, C.A., Y.F. Wiersma, and F. Huettmann (eds.). 2010. *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications*. Springer Verlag, New York, New York, USA. 314 p.

Franklin, J. 1995. Predictive vegetation mapping: Geographic modelling of biospatial patterns in relation to environmental gradients. *Prog. Phys. Geogr.* 19(4):474-499.

Gregory, R., D. Ohlson, and J. Arvai. 2006. Deconstructing adaptive management: criteria for applications to environmental management. *Ecological applications* : a publication of the Ecological Society of America 16(6):2411-2425.

Holling, C.S. (e.d.) 1978. Adaptive environmental assessment and management. John Wiley & Sons, New York, New York, USA.

Huettmann, F. 2005. Research and management viewpoint: databases and science-based management in the context of wildlife and habitat: toward a certified ISO standard for objective decision-making for the global community by using the internet. *Journal of Wildlife Management* 69:466 - 472.

Johnson, N.C., A.J. Malk, R.C. Szaro, and W.T. Sexton (eds.). 1999. Ecological Stewardship: A Common Reference for Ecosystem Management. Elsevier Science Ltd, Oxford, U.K.

Malone, T., J. Liang, and E.C. Packee. 2009. Cooperative Alaska Forest Inventory. Gen. Tech. Rep. PNW-GTR-785. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. 42 p.

Messier, C. and D. Kneeshaw. 1999. Thinking and acting differently for a sustainable management of the boreal forest. *Forestry Chronicle* 75: 929-938.

Ogden, A.E., and J. Innes. 2007. Incorporating climate change adaptation considerations into forest management planning in the boreal forest. *Int. For. Rev.* 9(3):713-733.

Ogden, A.E., and J.L. Innes. 2009. Application of Structured Decision Making to an Assessment of Climate Change Vulnerabilities and Adaptation Options for Sustainable Forest Management. *Ecology and Society* 14(1):11.

Olden, J.D., J.J. Lawler, and N.L. Poff. 2008. Machine learning methods without tears: A primer for ecologists. *Q. Rev. Biol.* 83(2):171-193.

Schwenk, W.S., T.M. Donovan, W.S. Keeton, and J.S. Nunery. 2012. Carbon storage, timber production, and biodiversity: comparing ecosystem services with multi-criteria decision analysis. *Ecological applications* : a publication of the Ecological Society of America 22(5):1612-1627.

Sokal, R., and N. Oden. 1978. Spatial autocorrelation in biology 1. Methodology. *Biological Journal of the Linnean Society* 10(2):199-228.

Sulak, A., and L. Huntsinger. 2012. Perceptions of Forest Health among Stakeholders in an Adaptive Management Project in the Sierra Nevada of California. *J. For.* 110(6):312-317.

Team, R.D.C. 2010. R: A language and environment for statistical computing. in R Foundation for Statistical Computing, Vienna, Austria.

Wagner, H.H., and M.J. Fortin. 2005. Spatial analysis of landscapes: Concepts and statistics. *Ecology* 86(8):1975-1987.

Walters, C. 1986. Adaptive management of renewable resource. Macmillan Publishing, New York, New York, USA.

Walters, C.J., and C.S. Holling. 1990. Large-Scale Management Experiments and Learning by Doing. *Ecology* 71(6):2060-2068.

Zhou, M., J.J. Liang, and J. Buongiorno. 2008. Adaptive versus fixed policies for economic or ecological objectives in forest management. *For. Ecol. Manage.* 254(2):178-187.

Table 5.1: Open Access spatial data for the boreal forest of Alaska

Open Access Data	Source
Bonanza Creek LTER	<a href="http://www.lter.uaf.edu/data">www.lter.uaf.edu/data</a>
Scenarios Network for Alaska Planning (SNAP)	<a href="http://www.snap.uaf.edu">www.snap.uaf.edu</a>
Alaska Geospatial Data Clearinghouse (AGDC)	<a href="http://www.agdc.usgs.gov/agdc.html">www.agdc.usgs.gov/agdc.html</a>
FSGeodata Clearinghouse	<a href="http://www.fsgeodata.fs.fed.us">www.fsgeodata.fs.fed.us</a>
Geographic Information Network of Alaska (GINA)	<a href="http://www.gina.alaska.edu/data">www.gina.alaska.edu/data</a>
Alaska Interagency Coordination Center	<a href="http://www.afsmaps.blm.gov">www.afsmaps.blm.gov</a>
Alaska Gap Analysis Project (AK-GAP)	<a href="http://www.aknhp.uaa.alaska.edu">www.aknhp.uaa.alaska.edu</a>
Alaska Mapped Statewide Digital Mapping Initiative (SDMI)	<a href="http://www.alaskamapped.org/sdmi">www.alaskamapped.org/sdmi</a>
USGS Landfire Data Distribution Site	<a href="http://www.landfire.cr.usgs.gov">www.landfire.cr.usgs.gov</a>